Control of binding and movement of fibroblast growth factors by heparan sulfate in extracellular matrix

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

By

Changye Sun

September 2015

Declaration

I declare that the work in this dissertation was carried out in accordance with the regulations of the University of Liverpool. The work described is original and has not been submitted for any other degree. All aspects of the experimental design and planning for the study were conducted by me in conjunction with my supervisors, Professor D Fernig, Dr R Lévy and Dr E Yates. The experimental work in this dissertation has been undertaken by me, with specific contributions that I have indicated below and in the text.

Yong Li, Sarah Taylor, Dr Xianqing Mao and Dr Mark Wilkinson assisted me in the production of FGF proteins. Dr Mal Horsburgh supported ideas for DNA cloning. Dr Marco Marcello provided expert guidance in the imaging data collection. Dr David Mason helped me to analyse the acquired imaging data.

Any views in this thesis are those of the author and in no way represent those of the University of Liverpool. This thesis has not been presented to any other university for examination in the United Kingdom or overseas.

Abstract

In animal development and tissue repair, gradients of extracellular signalling proteins, such as fibroblast growth factors (FGFs), are formed in the extracellular matrix (ECM) to conduct cell-cell communication. It has been argued that the heparan sulfate (HS) chains of proteoglycans are important in controlling the diffusion of these signalling proteins and the paracrine FGFs are classic HS binding proteins. There are over 883 proteins that bind to HS in the ECM, yet the extent to which protein binding is specific and selective is not clear. Moreover, the HS chains and the ECM in general are often considered to be a hydrated gel and have no structure, beyond the individual components. It is difficult to reconcile the regulation of critical and complex events in cell communication by HS and ECM with this view.

In this thesis, the production of a number of recombinant FGFs is described. This led to the discovery that HaloTag, a mutated chloroalkane dehalogenase conventionally used as a means to chemically label fusion proteins, is also an excellent solubilisation tag for FGFs. Thus, HaloTag fusions allowed the production and purification of members of the FGF family not generally available. These FGFs were used to characterise the interactions of FGFs with HS *in vitro* and in cells.

The interactions of FGF10 and FGF20 with HS were characterised by differential scanning fluorimetry, using a library of model polysaccharides. The data show that in terms of their preferred binding structures, FGF10 is most like its closely related subfamily member, FGF7. Likewise, binding preferences of FGF20 are most similar to those of FGF9, which is in the same subfamily. These results support the idea that the specificity and selectivity of the interactions between FGFs and HS have been driven by the same selection pressures that led to the diversification of the FGF family and to their binding selectivity for isoforms of their receptor tyrosine kinase.

To study the interaction with HS in cells, fixed Rama 27 fibroblasts were used, so that cell biochemistry would not confound the results. Fluorescent Halo-FGF-1, -2, -6, -10 and -20 were produced by reacting the HaloTag with a chloroalkane-TMR ligand. The binding of the fluorescent FGFs to Rama 27 cell pericellular matrix was measured by confocal microscopy. This showed that the binding sites for these FGFs

were heterogeneously distributed and the number of the binding sites for each FGF was different. Diffusion of these FGFs was measured by fluorescence recovery after photobleaching (FRAP). This demonstrated that the diffusion speed and the relative proportion of mobile *versus* immobile FGF were different for each FGF. The data indicate that the HS in pericellular matrix is selective in its binding of FGFs and determining how they diffuse.

In conclusion, the pericellular matrix exhibits high binding selectivity for different FGFs, which is consistent with the *in vitro* characterisation of their interactions with model polysaccharides. Moreover, cells spatially segregate the different binding structures in HS. This indicates that pericellular matrix is highly organised over length scales equivalent to hundreds or more HS chains, suggesting that it is formed by the self-organisation of its constituents though their extracellular interactions and, for transmembrane components, intracellular interactions as well. Such structure would then lie at the heart of how ECM performs its varied functions, including the selective binding and transport of FGFs.

Acknowledgements

There are so many people to thank for helping and supporting me to finish my project and being friends, providing advice on life during my four years PhD study.

Firstly, I would like to express deepest gratitude to my supervisors: David Fernig, Raphaël Lévy and Ed Yates. You have not only provided support in the great ideas for my research but also taught me to develop myself. So, now I have more confidence on the research that I did in the past four years and that I can do in the future.

I am grateful to the people who I have worked with and who helped me in the biochemistry lab at the University of Liverpool. Without your help my work would be much more difficult and I would spend more time to finish my project.

I would also like to thank my friends. No matter how far we are from each other, you always try to find time to care about me. Thanks for always being around and speaking and listening to me. I need to specially give my thanks to Yong Li, who has already been the best friend for seven years! Some friends, Jenny, Lara, Dom, Quentin, Ed and Dan, are also so important for me because you give me more reasons to enjoy life.

Finally, I would like to thank my dad, Lishu Sun and mum, Zhongyun Wang. You both are always giving me all the support to chase what I want and what I like. Your continuous love and trust encouraged me to complete my thesis.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BSA: bovine serum albumin

cDNA: complementary DNA

CS: chondroitin sulfate

DMSO: dimethyl sulphoxide

DMEM: Dulbecco's modified Eagle medium

dp: degree of polymerization

DS: dermatan sulfate

DSF: differential scanning fluorimetry

DT_{50}: half maximal denaturation temperature

ECM: extracellular matrix

FCS: fetal calf serum

FGF: fibroblast growth factor

FGFR: fibroblast growth factor receptor

FRAP: fluorescence recovery after photobleaching

GAG: glycosaminoglycan

GlcA: glucuronic acid

GlcN: glucosamine

GlcNAc: N-acetylglucosamine

HA: hyaluronic acid or hyaluronan

HaloTag: a mutant of a bacterial haloalkane dehalogenase

Halo-FGF: FGF with an N-terminal HaloTag

HBS: heparin binding site

HS: heparan sulfate

HSPG: heparan sulfate proteoglycan

IdoA: L-iduronic acid

IPTG: isopropyl β-D-1-thiogalactopyranoside

LB: lysogeny broth

MAPK: mitogen-activated protein kinases

OD600: optical density at 600 nm

PAGE: polyacrylamide gel electrophoresis

PBS: phosphate-buffered saline

PCR: polymerase chain reaction

PDB: protein data bank

PFA: paraformaldehyde

Rama: rat mammary

SDS: sodium dodecyl sulphate

SUMO: Small ubiquitin-like modifier protein

TAE: Tris-acetate-EDTA

TEMED: N,N,N',N'-Tetramethylethylenediamine

Tm: Melting temperature

Tris: Tris (hydroxymethyl) methylamine

Tris-NTA: Tris-Ni2+ nitriloacetic acid

Tween 20: polyoxyethylene sorbitan monolaurate

Chapter 1 Introduction

1.1 Cell-cell communication in development, homeostasis and disease

A fertilised egg cell will develop into an organism, such as a monkey or a fly (Fig. 1.1). Organisms are constructed of multiple cells, but they have very different structures, which arise through embryonic development. Cell-cell communication is one of the most important mechanisms contributing to embryogenesis, playing a critical role in coordinating cell functions and directing cell activities. By talking to each other, the cells are regulated to divide, to migrate to specific positions in the three dimensional space of the organism and to differentiate into another type of cell. For example, epithelial cells can differentiate into mesenchymal cells through cell-cell and cell-extracellular matrix (ECM) contacts (Radisky, 2005, Lamouille et al., 2014). Similarly, in tissue repair and homeostasis, cell division, differentiation and migration are also regulated by cell to cell communication (Chen et al., 2014).

The language of cell communication is signalling molecules. These are mainly proteins, which are secreted into the extracellular space. When these signalling molecules are released by the source cells, they must diffuse to their target cells, where, through receptor systems, they elicit the appropriate biochemical response to transmit signals and the biological response ensues. Based on the distance of the target from the source cells, paracrine and endocrine mechanisms are two well-established mechanisms of communication. Paracrine signalling molecules are restricted to the extracellular space surrounding the source cell and activate the

vicinal target cells, while the endocrine signalling molecules, hormones, are able to move systemically and so target distant cells. Autocrine is a special case of paracrine, by which the source cells signal to themselves. In development and tissue repair, the paracrine mechanism is critical to direct specific tissue and organ formation, such as limb development.

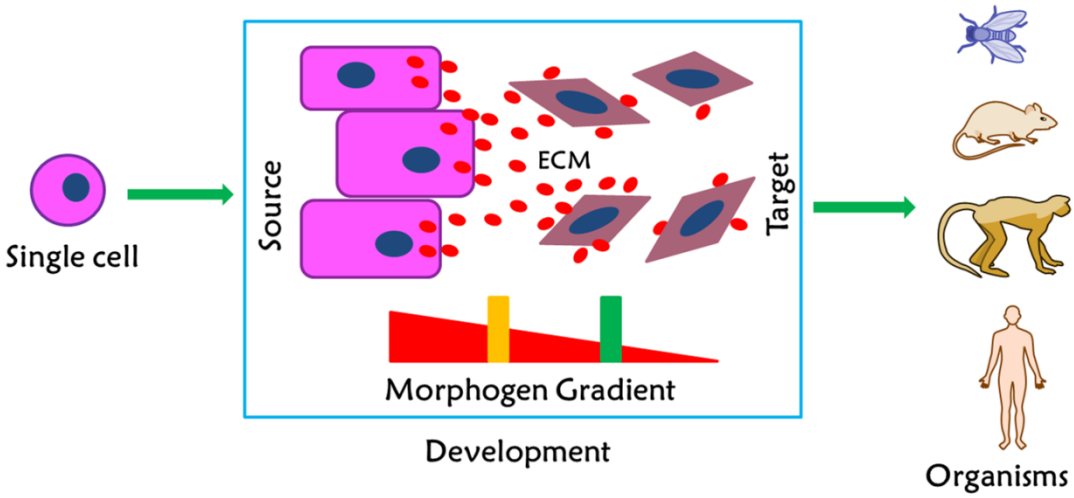


Figure 1.1 Cell-cell communication in development. A single fertilised cell can divide and differentiate into different types of cells, which form tissues and organs in the organism through the action of signalling molecules. One important feature is morphogenesis, whereby morphogen signalling molecules form gradients distributed in the ECM from source cell to the target cells that can recognise the different concentrations and consequently follow different fates. Thereafter, gradients of signalling molecules regulate many aspects of homeostasis.

1.2 Morphogen gradient formation and functions

Morphogens are the signalling molecules, secreted into the extracellular space by the source cells to form a concentration gradient among the surrounding target cells,

which induces concentration-dependent biological outcomes. In developmental biology, morphogen gradients play a critical role, because they spatially and temporally regulate cell fate and cell division, migration and differentiation, and so specify the body's structures are gradually developed (Sutherland et al., 1996, Gurdon and Bourillot, 2001). It is, therefore, important to know how morphogen gradients are formed, and how the cells create and maintain a particular morphogen gradient and concentration at specific positions (Fig 1.2 A) (Gurdon and Bourillot, 2001). Free diffusion is unlikely to allow appropriate gradients to form, because the concentration of morphogen will decay exponentially from the source. Consequently, over distances of a few cell diameters, the gradient will be extremely shallow and unlikely to be decoded (Qu et al., 2012, Zhou et al., 2012). Thus, for a gradient to form, the diffusion of the morphogen needs to be slowed down, which can be achieved by reversible binding to molecules in the ECM that lies between cells. Such molecules include the glycosaminoglycans (Belenkaya et al., 2004, Yan and Lin, 2009, Duchesne et al., 2012). Many classical morphogens, such as members of fibroblast growth factor (FGF), transforming growth factor-β (TGF-β), WNT and hedgehog families, have been identified in previous work and they share the property of binding to heparan sulfate (HS) (Gurdon and Bourillot, 2001).

In *Drosophila* wing development, the two important signalling molecules, Decapentaplegic (Dpp) and Wingless (Wg), form different morphogen gradients in the same ECM by means of different transport mechanisms (Kicheva et al., 2007). A number of different models have been proposed to regulate the transport of morphogens (Yan and Lin, 2009, Muller et al., 2013, Akiyama and Gibson, 2015), which fall into four main types (Figs 1.2 B1-B4). The free diffusion model does not require the signalling molecules to bind to any components of ECM. However, under

these conditions the signalling molecules could be degraded, since they are not protected, and concentration gradients would be very steep and only exist next to the source (Fig. 1.2 B2) (Muller et al., 2013, Akiyama and Gibson, 2015). In contrast, in the restricted diffusion model, the interaction of signalling molecules with ECM components, of which HS proteoglycans are the most important in this respect, shapes the morphogen gradient, because binding slows diffusion (Fig. 1.2 B1) (Lin, 2004, Kalinina et al., 2009, Yan and Lin, 2009, Makarenkova et al., 2009, Bokel and Brand, 2013, Akiyama and Gibson, 2015). The planar transcytosis model involves the transport from the source cells to the targets cells by repeated endocytosis and re-secretion by the intermediary cells (Fig. 1.2 B3) (Kicheva et al., 2007, Muller et al., 2013, Akiyama and Gibson, 2015). In this model, diffusion is slowed down by these cellular activities. Cytonemes have been proposed as another mechanism (Fig. 1.2 B4) (Akiyama and Gibson, 2015, Yan and Lin, 2009), although the HS proteoglycans on the cytonemes might play the role in transport of signalling molecules. Other models have been proposed, such as lipoprotein transport (morphogens diffuse after packing into lipoprotein particles), but the diffusion restricted by HS proteoglycans is thought to be the most common mechanism contributing to the formation of the morphogen gradients (Yan and Lin, 2009).

However, the mechanisms of morphogen gradient formation for many of these signalling molecules are still not very clear. Mutations of HS proteoglycans showed that the movement of Dpp-GFP in ECM was restricted by HS proteoglycans (Fig. 1.2 B2) (Belenkaya et al., 2004), while a kinetic study using fluorescence recovery after photobleaching (FRAP) suggested that dynamin-mediated endocytosis was required for the movement of Dpp-GFP (Fig. 1.2 B3) (Kicheva et al.). Thus, the data from living systems appear contradictory. Moreover, it is difficult in living systems to

probe one mechanism in detail and in isolation of other potential mechanisms. A fixed, so dead system might be a better means to understand the transport of morphogen molecules in the ECM. Such an approach would potentially also provide insight into the organisation and structure of the ECM.

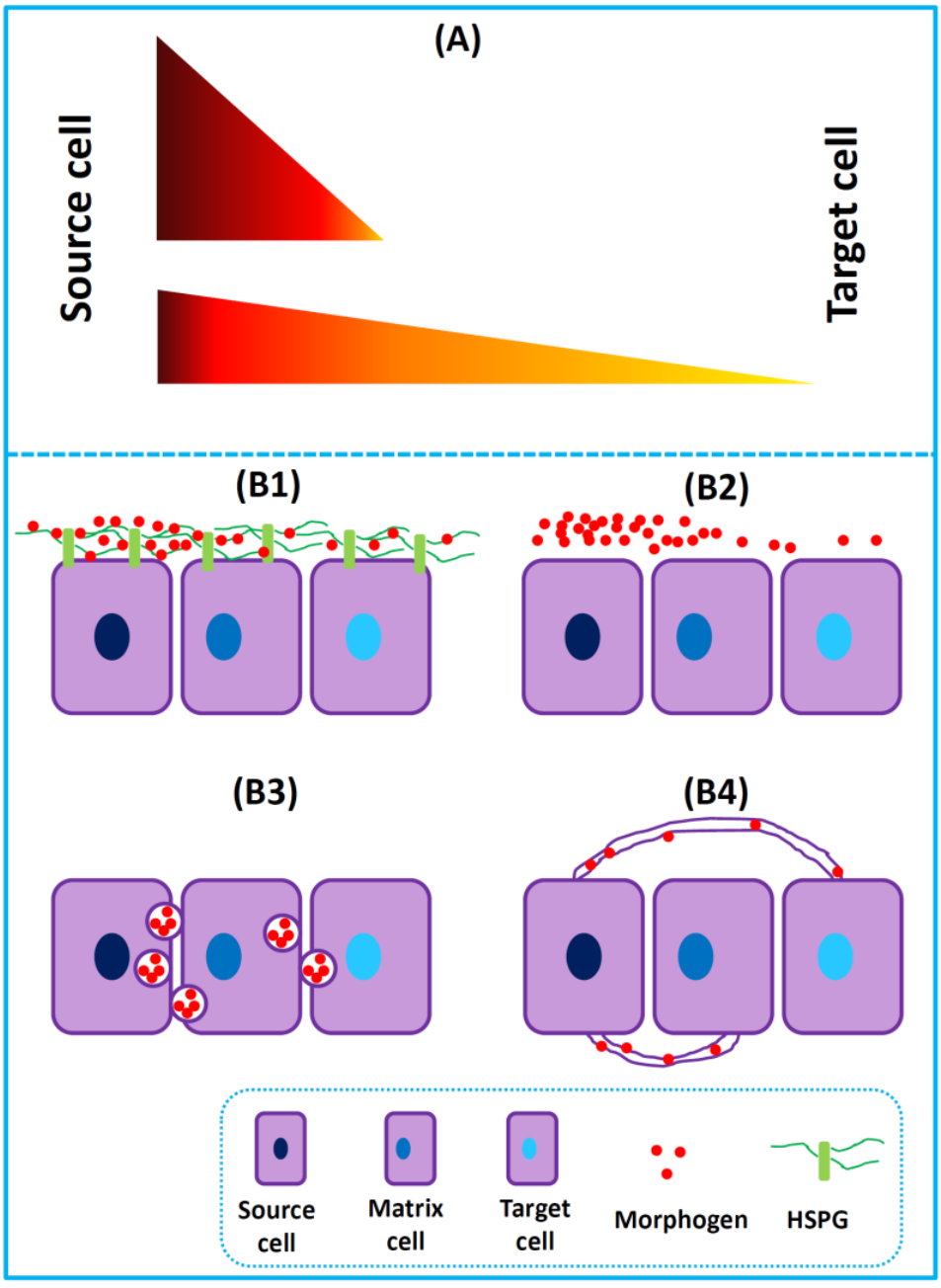


Figure 1.2 Main mechanisms proposed for morphogen transport and gradient formation. Morphogens (signalling molecules) secreted by source cells can form gradients of different shapes. The transport of these molecules is thought to be

regulated by different mechanisms, some of which are described in this figure. **(A):** Short-fat and long-slim morphogen gradients. **(B1-B4):** Transport mechanisms. **(B1):** Restricted diffusion by binding to HS proteoglycans. **(B2):** Free diffusion in the ECM. **(B3):** Planar transcytosis model, where the morphogens are transported by endocytosis and re-secretion. **(B4):** Diffusion of morphogens along the cytonemes to other cells/target cells.

1.3 ECM components and structure

There is clear evidence for interactions with the ECM controlling the movement of signalling molecules (Migliorini et al., 2015, Akiyama and Gibson, 2015, Bokel and Brand, 2013, Muller et al., 2013, Duchesne et al., 2012, Belenkaya et al., 2004), so, two important questions are: *what is in the ECM* and *is the ECM organised differently according to its functions*? The answer to the first question is becoming relatively clear. The ECM consists of three regions. The pericellular matrix, which includes proteins anchored on the cell surface, proteoglycans and restricted proteins (Fig. 1.3 grey region), extends around 35 nm from the cell surface. The gap area (Fig. 1.3 the middle area) between cells is classic ECM and the pericellular matrix of epithelial cells is separated from that of mesenchyme by basement membrane, which is an ECM, but named for its aspect under the light microscopy. ECM and basement membrane contain secreted proteins, which include fibrous proteins, such as collagens, and a range of glycoproteins, *e.g.*, fibronectin, laminin and proteoglycans (*e.g.*, perlecan) (Mouw et al., 2014, Frantz et al., 2010, Lingwood and Simons, 2010, Davies, 2001). These ECM components give rise to physical cell-cell connections, adhesion, signalling and provide the tissue with particular mechanical characteristics (Harisi and Jeney, 2015, Fan et al., 2014, Olczyk et al., 2014).

A large number of studies of tissue ECM have found that the components and the number of the components are very varied (Mori et al., 2014, Bonnans et al., 2014). Tissue staining reveals that the expression levels of collagen, laminin and fibronectin in subcutaneous adipose tissue and visceral adipose tissue are very different (Mori et al., 2014). Even in the same tissue, the components of ECM vary at different ages and for the normal and abnormal (disease) states, since the ECM is remodelled in development, tissue repair and often in diseases, which reflects changes (desired or otherwise) underlying biological functions (Lu et al., 2012, Frantz et al., 2010, Bonnans et al., 2014, Mori et al., 2014). For example, studies with rat skin revealed that the expression levels of collagen 1, laminin b1 and c1, and fibronectin 1 are strongly related to age (Mori et al., 2014, Quaglino et al., 1993). The structural analysis of HS in mice indicates that HS content, disaccharide composition and overall degree of sulfation are different for each individual tissue (Guimond et al., 2009, Ledin et al., 2004). The cancer cell microenvironment is also characterised by degradation and disorganisation of ECM (Lu et al., 2012). These types of observations indicate that it is likely that the ECM is an area with a highly dynamic structure rather than being a homogenous gel.

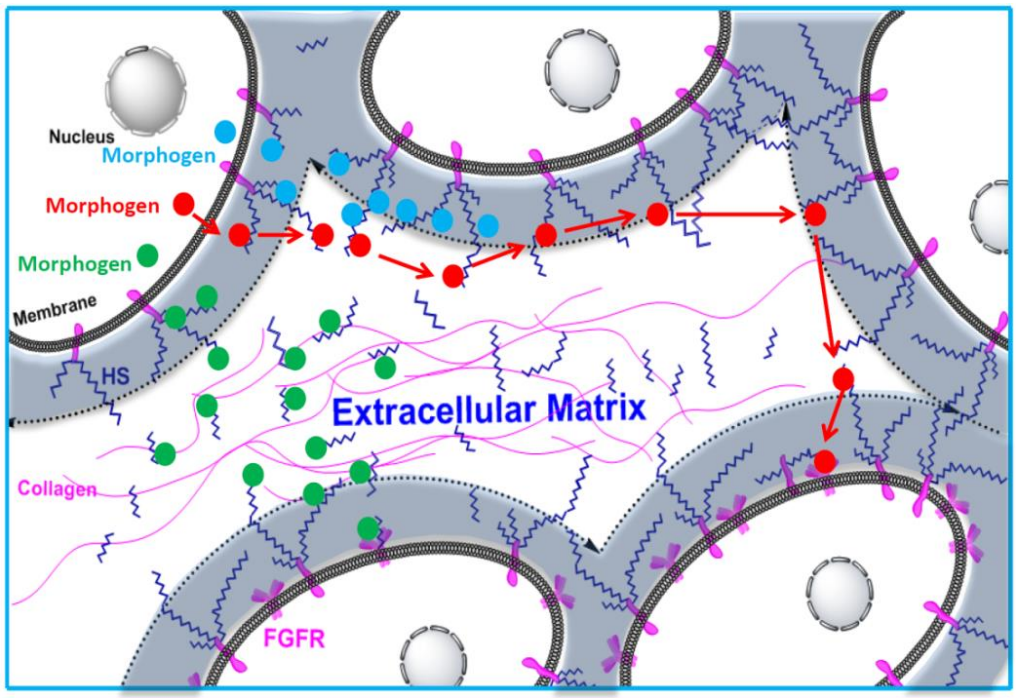


Figure 1.3 ECM components and structure, and diffusion of signalling proteins mediated by binding to HS chains. ECM includes the pericellular matrix (grey region) and the area between the cells (white region between cells). The pericellular matrix contains HS proteoglycans with core proteins inserted into the membrane, whereas the classic ECM contains secreted HS proteoglycans. Three different types of diffusion may be considered to arise from differential binding to HS chains. Weak binding to HS would lead to a fast diffusion (red dots), while strong binding might trap the HS binding proteins in the source cell matrix, which would then only release the HS binding proteins slowly (blue dots) or would require an external event, such as cleavage by heparanase to release the protein/protein-ligand complex (green dots).

1.4 HS proteoglycans

Proteoglycans are O-glycosylated proteins, such as perlecan, glypicans and syndecans (Taylor and Gallo, 2006, Yung and Chan, 2007). Some proteoglycans are inserted into the cell membrane and others are secreted into the ECM (Dreyfuss et al., 2009, Bernfield et al., 1999). These proteoglycans are expressed in the rough endoplasmic reticulum and then transported to the Golgi apparatus where the

glycosaminoglycan (GAG) chains are synthesised (Yanagishita and Hascall, 1992). The GAG chains are linear polysaccharides mainly consisting of repeating disaccharide units (Fig. 1.4 A) (Taylor and Gallo, 2006). The members of the GAG family are HS, chondroitin sulfate (CS), dermatan sulfate (DS), hyaluronan (HA) and keratan sulfate (KS) (Ori et al., 2008).

HS is made of repeating disaccharide units of glucuronic acid (GlcA) linked to N-acetylglucosamine (GlcNAc, Fig. 1.4 A). In the Golgi apparatus, the synthesis of HS GAG chains is started with the assembly of a tetrasaccharide linkage onto a serine residue of the core protein by four enzymes (Xyl transferase, Gal transferase I and II and GalA transferase); the repeat disaccharide units, $[\text{-GlcA-GlcNAc-}]_n$ (where n is ~25 to 100) are then added by the copolymerases EXT1 and EXT2 (Lin, 2004, Tumova et al., 2000, Dreyfuss et al., 2009). After the synthesis of the GAG chain, clusters of N-acetyl glucosamine are deacetylated and N-sulfate groups are added by the dual activity N-deacetylase-N-sulfotransferases (NDSTs) (Lin, 2004, Tumova et al., 2000, Dreyfuss et al., 2009). The subsequent modifications on N-sulfated glucosamine containing disaccharides or their neighbours: an epimerase converts glucuronic acid to iduronic acid (IdoA), which may then be 2-O sulfated and the glucosamine may be 6-O and 3-O sulfated (Lin, 2004, Tumova et al., 2000, Dreyfuss et al., 2009). Therefore, HS contains two different disaccharide units: GlcA-GlcNAc (without N-sulfation and with low sulfation groups on 6-O, 3-O and 2-O positions) and IdoA-GlcNAc (with N-sulfation and with more sulfation groups on 6-O, 3-O and 2-O positions).

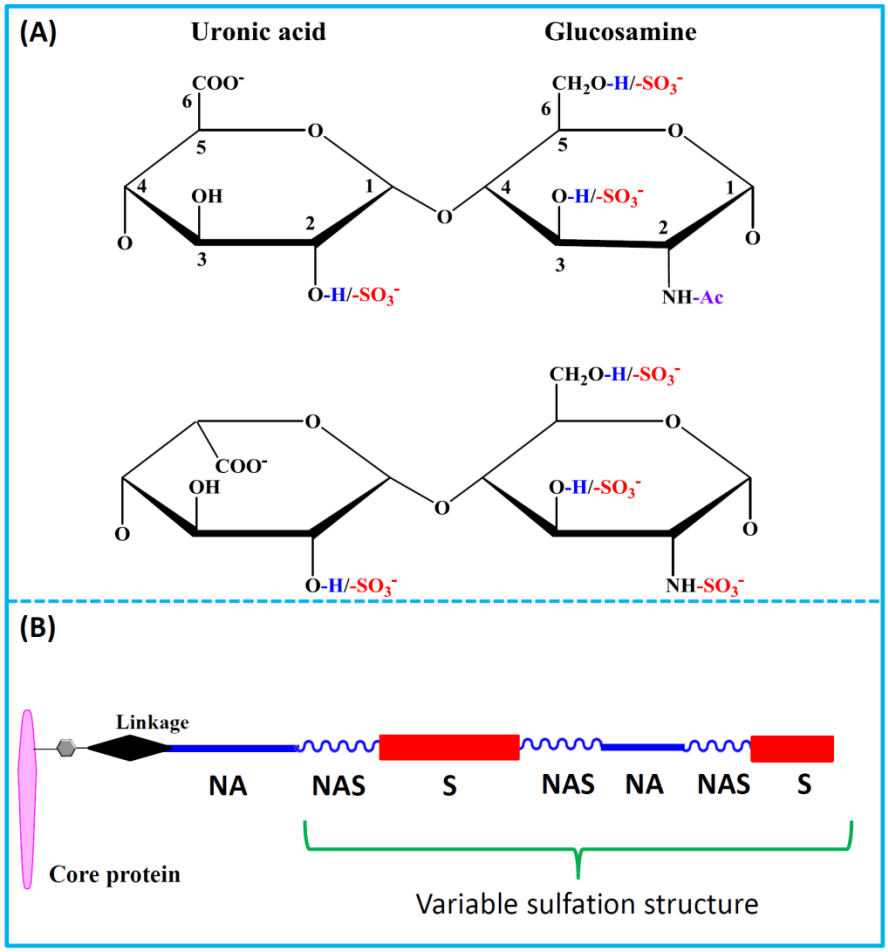


Figure 1.4 Structures of disaccharide units of HS and heparin. (A): Structure of disaccharide unit of heparin/HS. Top: The N-acetylglucosamine containing disaccharide. This is generally not or only slightly modified by O-sulfation (in red) and the uronic acid will always be glucuronic acid. Bottom: An iduronic acid containing disaccharide, which always contains an N-sulfated glucosamine (N-sulfate in red) and is often further modified by O-sulfation (red). **(B):** Structure of HS chains. The polysaccharide chain is covalently linked to a serine on the proteoglycan core protein. The sulfate groups are added by sulfotransferases after the GAG chain is polymerised. Owing to the hierachical dependence of the post polymerisation reactions and the mechanism of sulfation of discrete blocks of N-acetylglucosamines by N-deacetylase-N-sulfotransferase (NDST), the HS chain has a domain structure of alternating NA (GlcNAc), NAS (~ one disaccharide in two is N-sulfated) and NS (every glucosamine is N-sulfated) domains. Chain lengths vary from ~25 disaccharides to over 100. Heparin, a common experimental proxy for HS is ~30 disaccharides in length and can be considered to be a highly sulfated S domain.

Since NDST selectively acts on blocks of disaccharides, the modified HS chain has a domain structure of NA, NAS and NS domains (Fig. 1.4 B) (Dreyfuss et al., 2009, Ori et al., 2008, Connell and Lortat-Jacob, 2013). Differences in sulfation level of the NAS and NS domains provide the means for HS to bind with varying degrees of selectivity to over 883 proteins (Ori et al., 2008, Ori et al., 2011, Xu and Esko, 2014, Nunes, 2015); for example, FGF2 requires GlcNS-IdoA2S structures for binding, only found in S-domains, whereas antithrombin III binds a pentasaccharide containing both GlcNS and GlcNAc, which corresponds to a transition domain (Turnbull et al., 2001, Xu and Esko, 2014). Since the synthesis and modification reactions by the sulfotransferases do not go to completion, the length and level of sulfation of HS chains are also variable in different ECMs (Kirkpatrick and Selleck, 2007, Dreyfuss et al., 2009). Thus, different tissues have been found to contain different amount of HS with varying relative levels of the different sulfated disaccharides (Guimond et al., 2009, Ledin et al., 2004)

1.5 FGFs

1.5.1 Overview

Members of the FGF family regulate embryonic development and homeostasis, including the repair of tissues and organs and certain aspects of metabolism. In the past decades, 22 human FGFs have been identified. They are grouped into 7 subfamilies based on the analysis of their protein sequences (Fig. 1.5) (Ornitz and Itoh, 2001, Itoh and Ornitz, 2004). The FGF11 subfamily, known also as the fibroblast growth factor homology factors (FHF), exhibits high sequence and structural similarities with other FGFs, but the FHFs do not bind heparin or the FGF

receptor tyrosine kinase (FGFR) and are entirely intracellular (Beenken and Mohammadi, 2009). Consequently, FHFs are not always regarded as FGFs. The members of FGF 19 subfamily, which includes FGF19, FGF 21 and FGF 23, are endocrine factors, which act systemically and these three FGFs have very low or no heparin binding affinities (Itoh, 2007, Asada et al., 2009). The remaining five subfamilies (FGF1, FGF4, FGF7, FGF8 and FGF9) act locally. They are released into the ECM and their transport to their target cells is considered to be regulated by strong binding to HS in the ECM (Duchesne et al., 2012, Kirkpatrick and Selleck, 2007, Makarenkova et al., 2009). These FGFs signal by binding to the FGFR, but this also requires simultaneous engagement of the HS co-receptor (Zhu et al., 2010, Ornitz, 2000). Thus, HS regulates both the transport and the effector functions of these FGFs.

1.5.2 FGF evolution and functions

Itoh and Ornitz pointed out that FGF genes have not been found in unicellular organisms, such as *Escherichia coli* (*E. coli*) (Itoh and Ornitz, 2004). However, orthologous genes have been identified in all multicellular organisms, ranging from *Caenorhabditis elegans* (*C. elegans*) to *Homo sapiens*. Of the three FGFs in *Drosophila*, one called *branchless* induces branching of the trachea (Sutherland et al., 1996), while the *pyramus* and *thisbe* FGFs pattern the mesoderm of *Drosophila* embryos(Stathopoulos et al., 2004). Matus and coworkers isolated three FGF ligands (NvFGF8A, NvFGF8B, and NvFGF1A) from *Nematostella vectensis* and suggested their conserved roles in regulating gastrulation and development of neural tissue (Matus et al., 2007).

The evolution of FGFs is proposed to be divided into two major phases as the metazoan lineage extended to the chordate lineage (Itoh, 2007). In the first phase, the 2 or 3 FGF genes in primitive metazoa were expanded to 6 homologous genes by gene duplication (Itoh and Ornitz, 2004). The most studied extant primitive metazoan containing FGF genes is *C. elegans*, in which elg-17 and let-756 were identified (Ornitz and Itoh, 2001). Five of the six *fgfs* found in the tunicate *C. intestinalis* appear to share a common ancestor with vertebrate *fgfs*, which suggests the FGF genes of *C. intestinalis* lie between the first and second phases of the evolution of this growth factor family. Furthermore, another FGF gene of *C. intestinalis*, *Ci-fgf11/12/13/14*, and vertebrate *fgf11-fgf14* have identical intron-exon organization (Itoh and Ornitz, 2004). Therefore, the *fgfs* of the common ancestor of *C. intestinalis* and vertebrates are thought to have evolved into the FGF genes of vertebrates, such as *Homo sapiens* and Zebrafish, in the second phase (Itoh, 2007).

Basic FGF and acidic FGF, now known as FGF2 and FGF1, respectively, are the two original members of the FGF family (Fig. 1.5), and they can stimulate the proliferation of a variety of cells (Zhu et al., 1991). A large number of functions has since been found for each FGF. The knockout mouse model of FGF3 caused deafness (Tekin et al., 2007). FGF4 has the function of regulating the development of limb and cardiac valve leaflet formation, whereas FGF7 and FGF10, first known as keratinocyte growth factors (KGFs), are specifically expressed in mesenchyme where they signal to neighbouring epithelium (Beenken and Mohammadi, 2009). FGF6 was recognized as playing roles in myogenesis, which regulates muscle regeneration (Israeli et al., 2004). FGF8 has important roles in the development of brain, limb, eye and ear (Liu and Joyner, 2001, Lewandoski et al., 2000). Members of the FGF9 subfamily are involved in the stimulation of mesenchymal proliferation

(Beenken and Mohammadi, 2009). Many biological functions require several different FGF signaling pathways to work together. Taking the development of vertebrate limb as an example, mesenchymally expressed FGF10 elicits the development of the overlying apical ectodermal ridge, which subsequently secretes FGF8 to regulate the growth of the underlying mesoderm (Ornitz, 2000). Despite the substantial body of data on expression and spatial-temporal regulation of FGFs in development and homeostasis, there are still many questions about FGFs' functions to be answered.

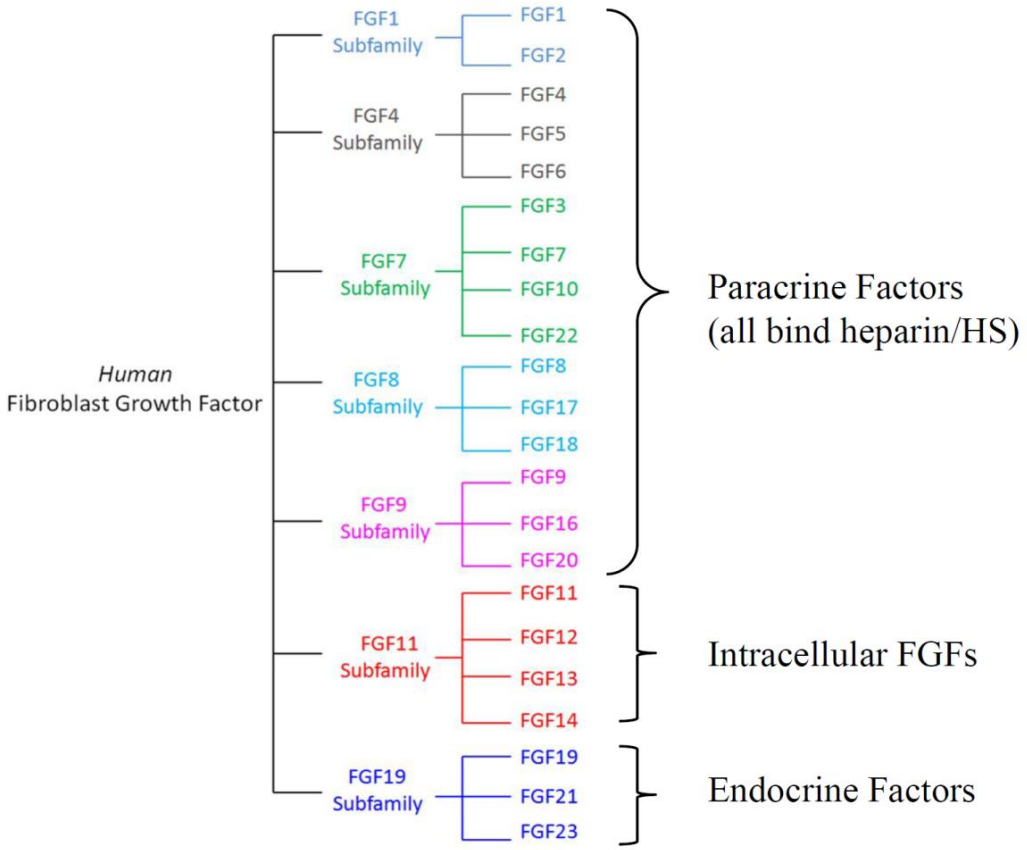


Figure 1.5 Family tree of human FGFs. Based on protein sequence alignment with the phylogenetic tree displayed by TreeView, the 22 FGFs are grouped into seven subfamilies: five paracrine factor subfamilies (FGF1 subfamily, FGF4 subfamily,

FGF7 subfamily, FGF8 subfamily and FGF9 subfamily), one endocrine factor subfamily (FGF19 subfamily) and one intracellular FGF subfamily (FGF11 subfamily).

1.5.3 FGF structures

Molecular weights of human FGFs range from 17 kDa to 34 kDa and the alignment of their protein sequences indicates that the core region of FGFs shares many conserved residues responsible for binding FGFR and HS, as well as the position of β-strands (Ornitz, 2000, Itoh and Ornitz, 2004). Thus, FGFs of *Homo sapiens* share more than 30 highly conserved amino acid residues (Fig. 1.6 A). Owing to the conservation of FGFs' primary structures in *Homo sapiens*, they have a similar three dimensional structure of 12 antiparallel β strands that form a β-trefoil fold (Figs 1.6 B-C).

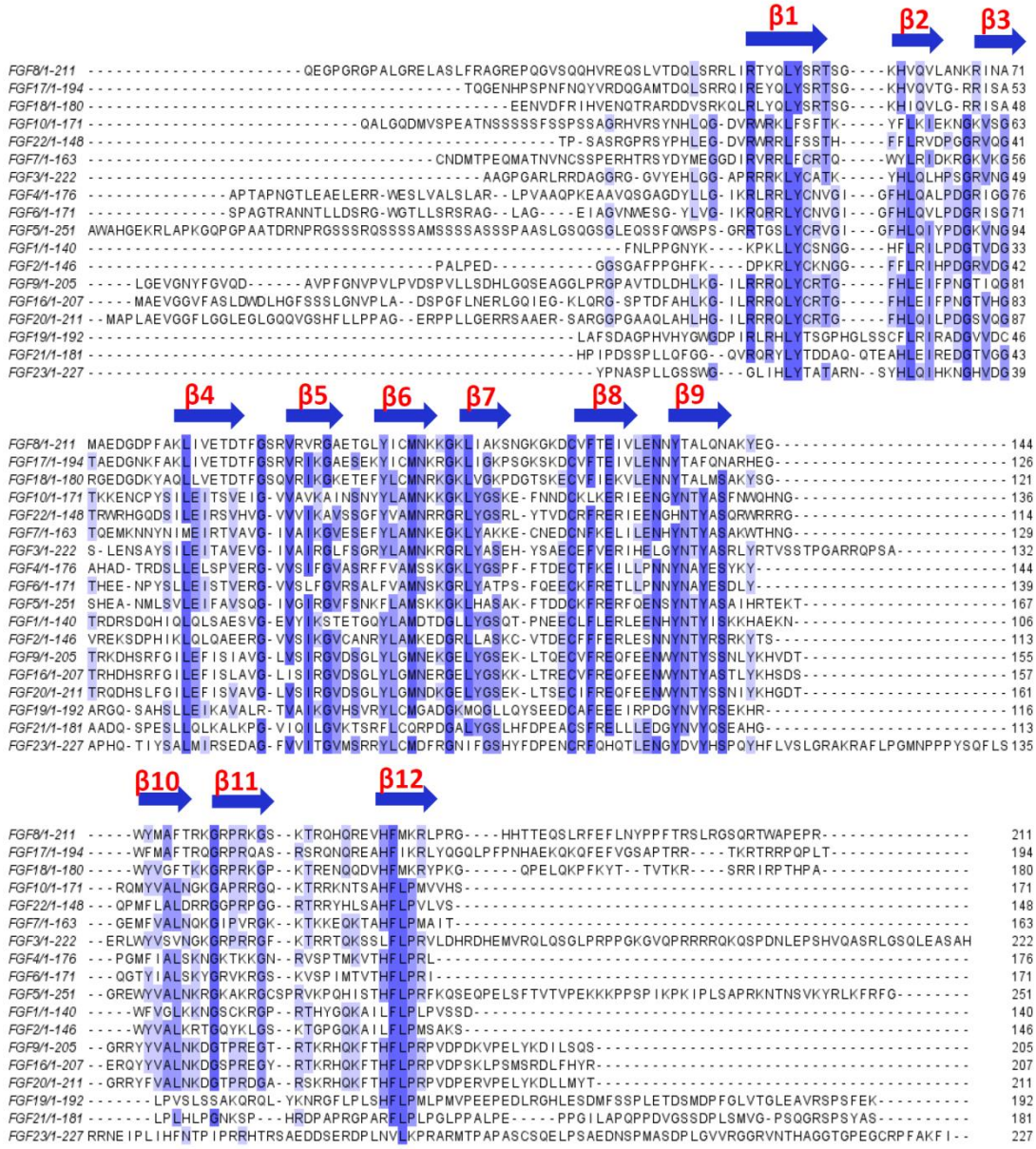


Figure 1.6 Sequence alignment of FGFs. The protein sequences of human paracrine FGFs and endocrine FGFs without the signal peptide were extracted from UniProt. These sequences were aligned by ClustalX 2.0.10 and the alignment was viewed with Jalview 2.6.1. The aligned residue positions with the most identical amino acids were colored with deep blue and residue positions with fewer identical amino acids were colored with light blue. It is important to note that this analysis examines identity only and not similarity. The blue arrows indicate the positions of the 12 β-strands.

The FGFs contain binding sites for heparin/HS (heparin/HS binding sites or HBS) and for the FGFR (Fig. 1.7 A). Canonical heparin/HS interacting amino acids of FGF1 and FGF2 locate at the β1-β2 loop and the region between the β10 strand and the β12 strand, which are rich in basic amino acids (arginine and lysine) (Figs 1.7 A-B) (Faham et al., 1996, Schlessinger et al., 2000, Pellegrini et al., 2000, Xu et al., 2012). Interestingly, this heparin-binding site is flexible such that the β10 and β11 strands are not well defined in the crystal structures (Fig. 1.7 & (Schlessinger et al., 2000)) and β11 is absent in the NMR structure (Moy et al., 1996); this may be important for the interaction with HS. FGFR binding sites are physically distinct and generally contained in the regions that include part of the N-terminus (around 10 amino acids before β1 (Olsen et al., 2006)), β1- β2 strands, β3 strand, β4-β5 strands, β6 strand, β7- β9 strand (Fig. 1.7 A and B) and the terminus of β12 strand. The primary sequences of these regions are very divergent among FGFs, which underlies the selectivity of different FGFRs for particular FGF ligands (Plotnikov et al., 2000, Mohammadi et al., 2005).

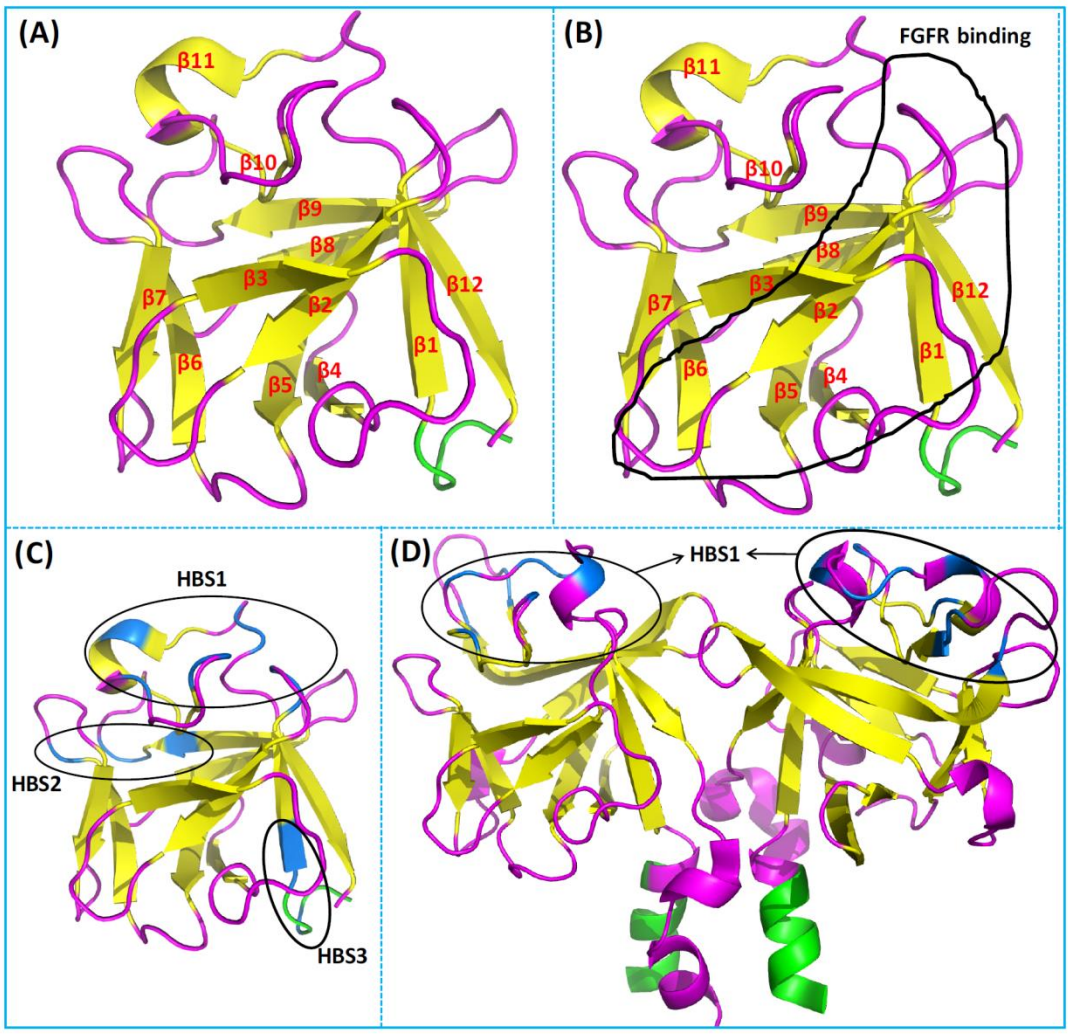


Figure 1.7 Structure of FGF2 and its interactions with FGFR and heparin/HS. **(A):** Structure of FGF2 (PDB: 1FQ9 ***(Schlessinger et al., 2000)***). **(B):** Surface of FGF2 that interacts with FGFR1 (from PDB: 1FQ9 ***(Schlessinger et al., 2000)***). **(C):** Heparin binding sites of FGF2 (PDB: 1FQ9 ***(Schlessinger et al., 2000)***) identified by a selective labelling approach ***(Ori et al., 2009)***. Three binding sites were recognised: the canonical binding site (HBS1), and two secondary and relatively weaker binding sites (HBS2 and HBS3). **(D):** HS binding site of FGF9 (PDB: 1G82 ***(Hecht et al., 2001)***). Only the conserved HBS1 was identified by selective labelling ***(Xu et al., 2012)***. Green is the N-terminus of the proteins. Grey is FGFR1. Magenta coloured proteins are FGFs (FGF2 in A, B and C, and FGF9 in D) with yellow coloured β strands. The residues in blue are the identified heparin binding sites of the FGFs.

1.5.4 Interaction of heparin/HS with FGFs

A great deal of work has been done to analyse the binding specificity of FGFs with heparin, heparin derivatives and HS using a wide variety of assays, *e.g.*, thermal stabilisation, heparin or FGF affinity chromatography (Turnbull et al., 1992, Maccarana et al., 1993, Jemth et al., 2002, Luo et al., 2006, Uniewicz et al., 2010, Xu et al., 2012). FGF1 and FGF2 are the two most studied FGFs in this respect. Originally, FGF2-affinity chromatography revealed that a tetradecasaccharide from HS (dp14, dp: degree of polymerisation) of 2-O-sulfated iduronate and N-sulfated glucosamine bound FGF2 with high affinity and the 6-O sulfate was not important (Turnbull et al., 1992). Maccarana *et al.* also found that the 2-O-sulfate of a smaller heparin dp5 was essential for FGF2 interaction, whereas 6-O sulfate groups were not required (Maccarana et al., 1993). In terms of FGF2 activity in mitogenesis assays, longer heparan sulfate oligosaccharides of at least 12 sugars (dp12) were required (Walker et al., 1994). Later, the crystal structure of heparin-FGF2 complexes demonstrated that both tetra- and hexasaccharides with 2-O-sulfate and N-sulfate groups can bind to FGF2 (Faham et al., 1996). A study with N-sulfated heparin oligosaccharide libraries suggested that FGF2 required either an octasaccharide with at least one 2-O-sulfate for binding. Alternatively, if the oligosaccharide lacked 2-O-sulfate groups, three adjacent 6-O-sulfated disaccharides were necessary (Jemth et al., 2002). Recently, the protein-ligand interaction study by a thermal shift assay also indicated that N-sulfate and 2-O-sulfate are important for FGF2 binding, but not 6-O-sulfate (Table 1.1) (Uniewicz et al., 2010). Studies using more heavily sulfated heparin-derived oligosaccharides have shown that heparin dp4 is the minimal length for FGF2 binding and activity and heparin dp8 can bind to both FGF2 and FGFR in the ternary crystal structure (Guglieri et al., 2008, Uniewicz et al., 2010, Schlessinger

et al., 2000). These differences may simply reflect the fact that heparin is highly sulfated and to obtain sulfates in the appropriate spatial orientation in HS requires longer structures than with heparin.

Table 1.1 Summary of previous findings on the binding selectivity of FGFs to heparin, heparin derivatives and heparin oligosaccharides (Uniewicz et al., 2010, Xu et al., 2012). Differential scanning fluorimentry (DSF) was used to measure the stabilisation effect of binding of FGFs to heparin and a library of model polysaccharides. The stabilisation effect is reported as an increase in the melting temperature (Tm) of the FGF. dp is degree of polymerisation.

Subfamily	Name	Heparin effect on Tm	Sulfation preference	dp length
FGF1	FGF1	> 22°C	NS ≈ 2S ≈ 6S	dp2 ~ dp10
	FGF2	> 22°C	NS, 2S > 6S	dp4 ~ dp10
FGF7	FGF7	~ 7°C	NS ≈ 2S ≈ 6S	dp4 ~ dp8
FGF8	FGF18	> 15°C	NS ≈ 6S > 2S	dp6 ~ dp10
FGF9	FGF9	~ 20°C	NS ≈ 6S > 2S	dp4 ~ dp10

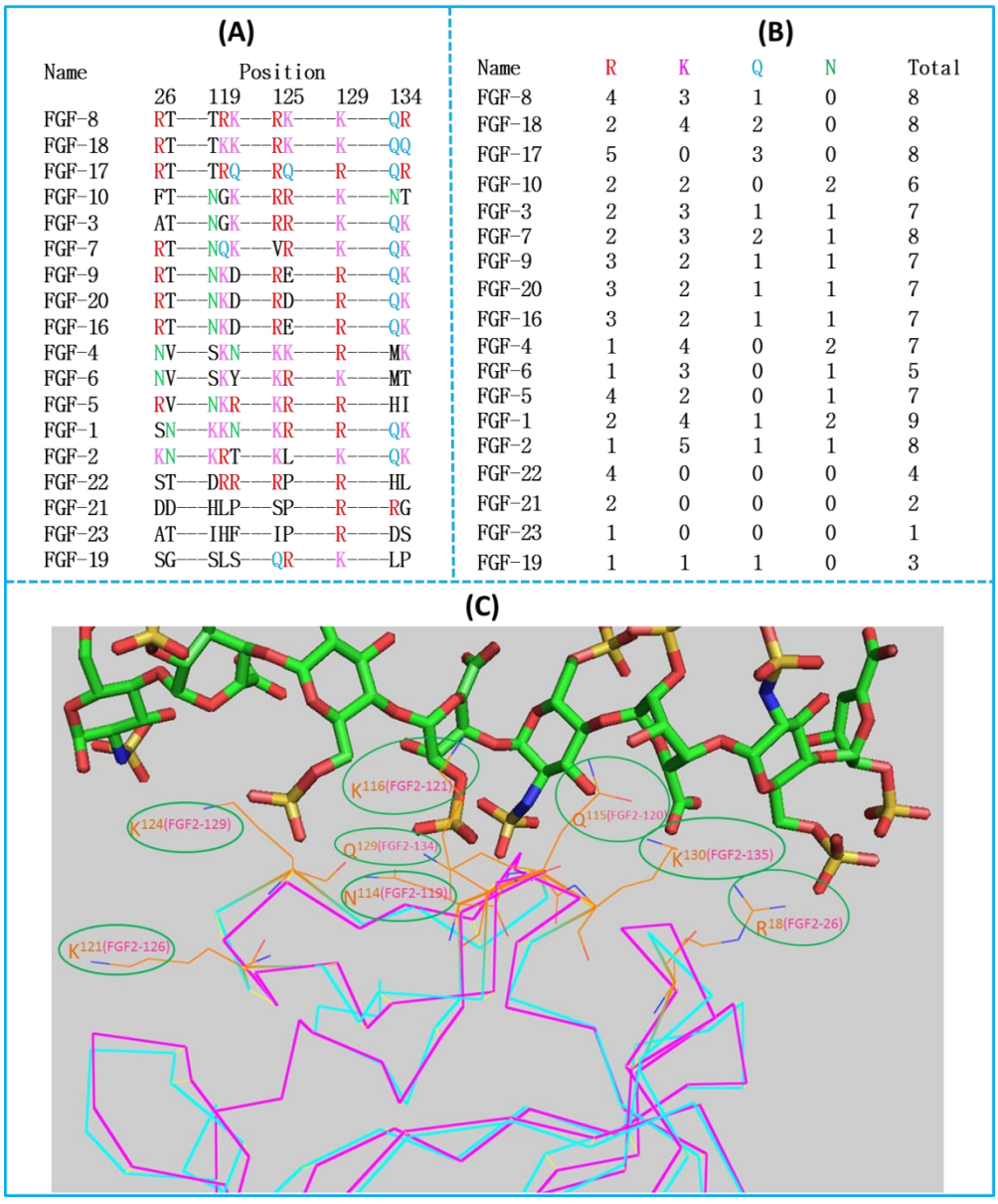
Both the N-sulfated oligosaccharide libraries and a library of chemically modified heparin derivative illustrate that FGF1 can bind to structures containing any two of 6-O-sulfate, 2-O-sulfate and N-sulfate (Jemth et al., 2002, Uniewicz et al., 2010). The selective binding of heparin octasaccharides using immobilised FGF chromatography demonstrated that FGF1, FGF4 and FGF8 can bind to N-sulfated oligosaccharides with 2-O-sulfate or 6-O-sulfate groups, but that more than two of these sulfate groups were required for optimal binding (Kreuger et al., 2005). The thermal stability assay suggests that FGF7 has a preference for 6-O sulfate and N-sulfate groups in heparin (Table 1.1) (Xu et al., 2012), and Luo *et al.* found

octasaccharides with 7 or 8 sulfate groups strongly bind to FGF7 (Luo et al., 2006). FGF9, which is a dimer (Hecht et al., 2001), prefers to bind more heavily charged heparin derivatives, which have only had one sulfate group (2-O-/6-O-/N- sulfate) removed (Xu et al., 2012).

In terms of the minimal length of saccharide binding structures, FGF1, FGF2 and FGF7 can bind to heparin dp4, while FGF9 and FGF18 require dp6 to bind (Uniewicz et al., 2010, Xu et al., 2012). It was found that FGF1, FGF2, FGF7, FGF9 and FGF18 can be efficiently stabilised by HS chains (Uniewicz et al., 2010, Asada et al., 2009). However, HS contains fewer sulfate groups and is more diverse in structure, thus a longer HS structure may be required to bind to these FGFs than with heparin (Walker et al., 1994, Turnbull et al., 1992). However, taken together these data suggest that FGFs bind to heparin and HS with a degree of selectivity. By binding to HS, FGFs can also be stabilised, protected from degradation by proteases and stored in the HS-rich pericellular matrix (Nur-E-Kamal et al., 2008, Uniewicz et al., 2010, Vlodavsky et al., 1987). Since HS from different cells and tissues possesses differences in sulfation patterns and domain structures (Ledin et al., 2004, Guimond et al., 2009), it seems reasonable to suggest that this may allow the regulation of the binding and the diffusion of FGFs in the ECM.

Previous work identified secondary binding sites for the polysaccharide on some FGFs (Xu et al., 2012, Ori et al., 2009). FGF1, FGF2 have three distinct heparin binding sites, while FGF9 would appear to have only the canonical binding site (Figs 1.7 B-C) (Xu et al., 2012, Ori et al., 2009). FGF7 also has an extended canonical binding site and a secondary binding site, which is on the side of canonical binding site, running orthogonal to this (Xu et al., 2012). An extended canonical heparin

binding site (HBS) was also found on the surface of FGF18, which also possesses one secondary binding site found on FGF1 and FGF2 (Xu et al., 2012). The canonical binding site is the most conserved binding site in the FGF family and is of highest affinity, so it is the primary driver of the interaction of the paracrine FGFs with HS (Ori et al., 2009, Makarenkova et al., 2009).



(A)

Name	Position				
	26	119	125	129	134
FGF-8	RT	TRK	RK	K	QR
FGF-18	RT	TKK	RK	K	QQ
FGF-17	RT	TRQ	RQ	R	QR
FGF-10	FT	NGK	RR	K	NT
FGF-3	AT	NGK	RR	K	QK
FGF-7	RT	NQK	VR	K	QK
FGF-9	RT	NKD	RE	R	QK
FGF-20	RT	NKD	RD	R	QK
FGF-16	RT	NKD	RE	R	QK
FGF-4	NV	SKN	KK	R	MK
FGF-6	NV	SKY	KR	K	MT
FGF-5	RV	NKR	KR	R	HI
FGF-1	SN	KKN	KR	R	QK
FGF-2	KN	KRT	KL	K	QK
FGF-22	ST	DRR	RP	R	HL
FGF-21	DD	HLP	SP	R	RG
FGF-23	AT	IHF	IP	R	DS
FGF-19	SG	SLS	QR	K	LP

(B)

Name	R	K	Q	N	Total
FGF-8	4	3	1	0	8
FGF-18	2	4	2	0	8
FGF-17	5	0	3	0	8
FGF-10	2	2	0	2	6
FGF-3	2	3	1	1	7
FGF-7	2	3	2	1	8
FGF-9	3	2	1	1	7
FGF-20	3	2	1	1	7
FGF-16	3	2	1	1	7
FGF-4	1	4	0	2	7
FGF-6	1	3	0	1	5
FGF-5	4	2	0	1	7
FGF-1	2	4	1	2	9
FGF-2	1	5	1	1	8
FGF-22	4	0	0	0	4
FGF-21	2	0	0	0	2
FGF-23	1	0	0	0	1
FGF-19	1	1	1	0	3

(C)

Figure 1.8 The canonical heparin/HS binding site of paracrine and endocrine FGFs. The ten residues in the canonical HBS1 were extracted from the aligned sequences (Fig. 1.6 A). These sequences were aligned again to test the conservation

across subfamilies. **(A):** Sequence alignment of the 10 residues in HBS1 for FGFs. **(B):** The numbers of residues with side chains known to interact with the polysaccharide in HBS1: arginine (R), lysine (K), asparagine (N), glutamine (Q). **(C):** Structure alignment of HBS1 of FGF2 (PDB: 1FQ9 (Schlessinger et al., 2000), magenta string) and FGF7 (PDB: 1QQK (Ye et al., 2001), blue string). The eight labelled residues from FGF7 correspond to the heparin binding residues of FGF2.

Ten amino acids in the core of the canonical HBS of FGF1 and FGF2 contribute to the interaction (Faham et al., 1996, Schlessinger et al., 2000, Pellegrini et al., 2000). Taking FGF2 as an example, the ten residues are K^{26}, N^{27} (β1-β2 loop), K^{119}, R^{120}, T^{121}, K^{125}, L^{126}, K^{129}, Q^{134} and K^{135} (the region between the β10 strand and the β12 strand), but the side chain of L^{126} is not a binding site (Fig. 1.8 A) (Faham et al., 1996, Schlessinger et al., 2000). The amino acids at the same position with the ten selected amino acids from FGF1 and FGF2 were extracted from the aligned sequences of the 18 paracrine and endocrine FGFs (FGF1-FGF10 and FGF16-FGF23). These sequences were then aligned by ClustalX 2.0.10 to compare the collinear part of the binding sequences. Arginine (R), lysine (K), glutamine (Q) and asparagine (N) were classed as heparin binding amino acids (Figs 1.8 A-B) (Faham et al., 1996, Schlessinger et al., 2000, Pellegrini et al., 2000), though this ignores contributions from the amide of the peptide backbone. The crystal structure of FGF7 was aligned to that of FGF2 as an example, and the ten residues on FGF7 were identified, which illustrates that these are accessible to heparin/HS (Fig. 1.8 C). As described, the members in the FGF19 subfamily have the lowest heparin/HS binding affinity and most of the recognized amino acids do not bind to the charged groups in heparin (Figs 1.8 A-B) (Asada et al., 2009, Xu et al., 2012, Ornitz, 2000). The other FGFs have similar number of heparin binding amino acids in the canonical heparin

binding region and they possess good heparin binding affinities based on heparin-affinity chromatography (Asada et al., 2009, Sun et al., 2015). However, in those cases where it has been measured, the FGFs prefer different structures in heparin polysaccharide, which suggests the binding of FGFs to HS is not simply primarily by protein surface charge (Figs 1.8 A-B) (Asada et al., 2009, Sun et al., 2015).

1.6 FGFRs and FGFR binding specificity of FGFs

FGFRs, anchored on the membrane, are the key to transferring induced signals into the cell, which direct the target cell activities, such as cell proliferation, differentiation and migration (Turner and Grose, 2010, Beenken and Mohammadi, 2009, Ornitz, 2000). Five different FGFRs (FGFR1-4 and FGFRL1) and many of their alternative spliced isoforms have been found to bind with FGFs and induce different signalling pathways (Powers et al., 2000, Itoh and Ornitz, 2011, Wiedemann and Trueb, 2000). FGFR1-4 possess three extracellular immunoglobulin-like loops, I, II and III (often termed D1, D2 and D3), a transmembrane linker and a cytoplasmic kinase domain (Fig. 1.9) (Powers et al., 2000, Beenken and Mohammadi, 2009, Goetz and Mohammadi, 2013). FGFRL1 differs in that its intracellular domain lacks a tyrosine kinase (Wiedemann and Trueb, 2000, Kim et al., 2001, Sleeman et al., 2001). The FGF ligand binds to surfaces in D2, the D2-D3 linker and D3 (Figs 1.9 A and B). In FGFR1, FGFR2 and FGFR3, half of D3 is encoded by alternative exons, denoted ‘a’, ‘b’ and ‘c’. The latter two give rise to the ‘b’ and ‘c’ isoforms of the transmembrane receptor, which impart additional ligand selectivity (Fig. 1.9 A and Table 3.2) (Ornitz et al., 1996, Zhang et al., 2006b).

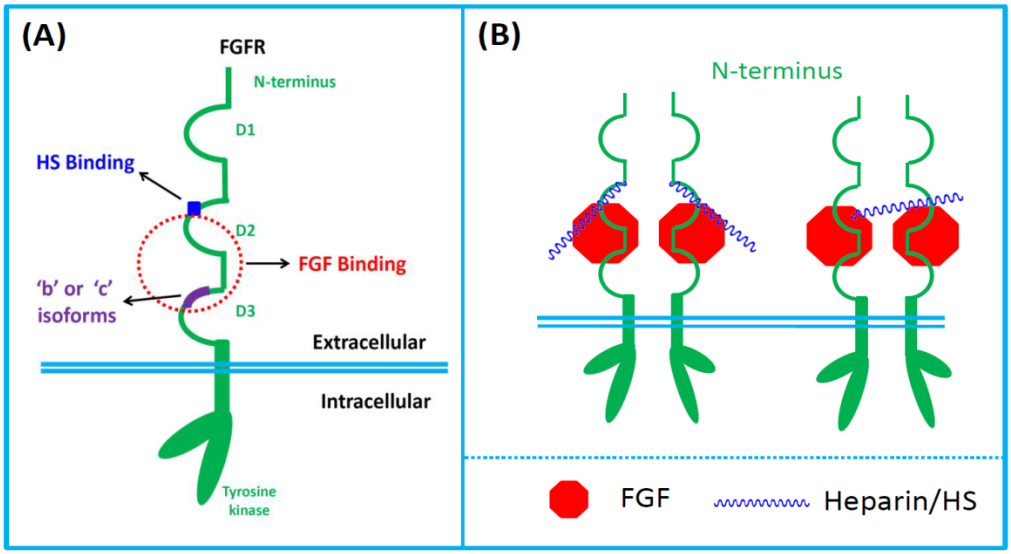


Figure 1.9 Structure of FGFR and FGF-FGFR-HS complex. (A): Structure of FGFR. The extracellular part of FGFR1-4 contains 3 immunoglobulin-like domains (D1-D3), a transmembrane helix and the intracellular tyrosine kinase. The HS binding site of FGFR is located at the D2 domain and FGF ligands interact with D2 and D3 domains and their linker. The splicing of D3 domain gives rise to 'b' and 'c' FGFR isoforms, which regulate the specificity of interaction of FGF and FGFR (Table 3.2). (B): FGF-FGFR-HS complex models. The two brief well-known models show the binding status of FGF-FGFR-HS and the stoichiometric ratio of FGF, FGFR and HS in the binding complex (Schlessinger et al., 2000, Pellegrini et al., 2000). The FGF-FGFR complex always happens in couples to induce the dimerization of receptors.

The FGFRs have varying degrees of selectivity for different FGFs, and the selectivity is most conserved between FGFs in the same subfamily (Zhang et al., 2006b, Xu et al., 2013). FGF1 was recognised as a universal ligand for all the FGF receptors, while FGF2 and members of the FGF4 subfamily prefer to interact with FGFR 1c (Zhang et al.). For the 'c' isoform, the preference is FGFR1c > FGFR 2c and FGFR 3c, though the FGF4 subfamily is clearly distinguished from the FGF1 subfamily in terms of their selectivity for FGFR 1b, which they do not bind, in contrast to FGF1 and FGF2 (Zhang et al., 2006b, Ornitz et al., 1996). Members of

FGF8 and FGF9 subfamilies preferentially bind to FGFR 3c (FGFR 3c > FGFR 2c and 1c), while members of the FGF7 subfamily mainly bind FGFR 2b and 1b (Zhang et al., 2006b). The alternative splicing of the D3 domain of the same receptor provides a means of imparting directionality and reciprocity to FGFs mediating cell communication, that is between two cells each acting as source/target for the other (Turner and Grose, 2010, Goetz and Mohammadi, 2013, Yu and Ornitz, 2001). For example, in limb development, FGFR2 IIIb is restrictively expressed by epithelial cells, while FGFR2 IIIc is mainly expressed by mesenchymal cells (Orrurtreger et al., 1993, Johnson et al., 1991). In contrast, the ligands binding to FGFR 2b, FGF7 and FGF10, are expressed and secreted by mesenchymal cells and FGFR 2c ligands (FGF8 and FGF9) are secreted by epithelial cells (Yu and Ornitz, 2001, Ornitz and Marie, 2002, Goetz and Mohammadi, 2013). In addition, the tyrosine kinase FGFRs were also found to bind HS, which contributes to a stable ternary FGF-FGFR-HS structure (Fig. 1.9 B) (Schlessinger et al., 2000, Powell et al., 2002, McKeehan et al., 1998). The most well-known models are 2:2:2 (symmetric) and 2:2:1 (asymmetric) FGF-FGFR-HS complex (Fig. 1.9 B). HS/heparin can facilitate the stability of FGF-FGFR complex to bring about some special biological functions, such as mitogenic activity (Mohammadi et al., 2005). In the asymmetric model, only one HS/heparin chain binds to the two FGFs, and this engages two FGF receptors (Pellegrini et al., 2000). The interaction of FGFs with their receptors induces the dimerization of receptors (Fig. 1.9 B), which regulates the phosphorylation of the intracellular kinase domain to induce intracellular signalling (Section 1.7).

1.7 FGF signalling pathways

The binding of the FGF ligand to its receptor causes the FGFR to dimerise. This in turn enables phosphorylation of tyrosine residues in the kinase activation loop and then of tyrosines that are docking sites for signalling proteins (Goetz and Mohammadi, 2013). These phosphorylations cause the activation of many intracellular signalling pathways, *e.g.*, RAS-RAF-MAPK, PI3K-AKT and phospholipase-IP3 (Fig. 1.10), which regulate cell fate and specific cell activities (Fig. 1.10) (Turner and Grose, 2010, Dorey and Amaya, 2010). Previous studies suggest HS/heparin is required for most, but not all signalling (Izvolsky et al., 2003, Delehedde et al., 2000, Delehedde et al., 2002b). FGF signalling can be negatively regulated by internalisation and degradation of the ligand-receptor complex, as well as by transmembrane regulators, such as FGFRL1 and intracellular ones, *e.g.*, MAPK phosphatase 3, sprouty and spred proteins (Turner and Grose, 2010, Casci et al., 1999, Eblaghie et al., 2003) . Since there is a great diversity of FGF ligands, FGFR isoforms, HS structure and feedback loops, the understanding of FGF signalling is still far from complete, though the link between the activation of the RAS-RAF-MAPK pathway and the stimulation of cell division is well established, as least in cultured cells (Dorey and Amaya, 2010).

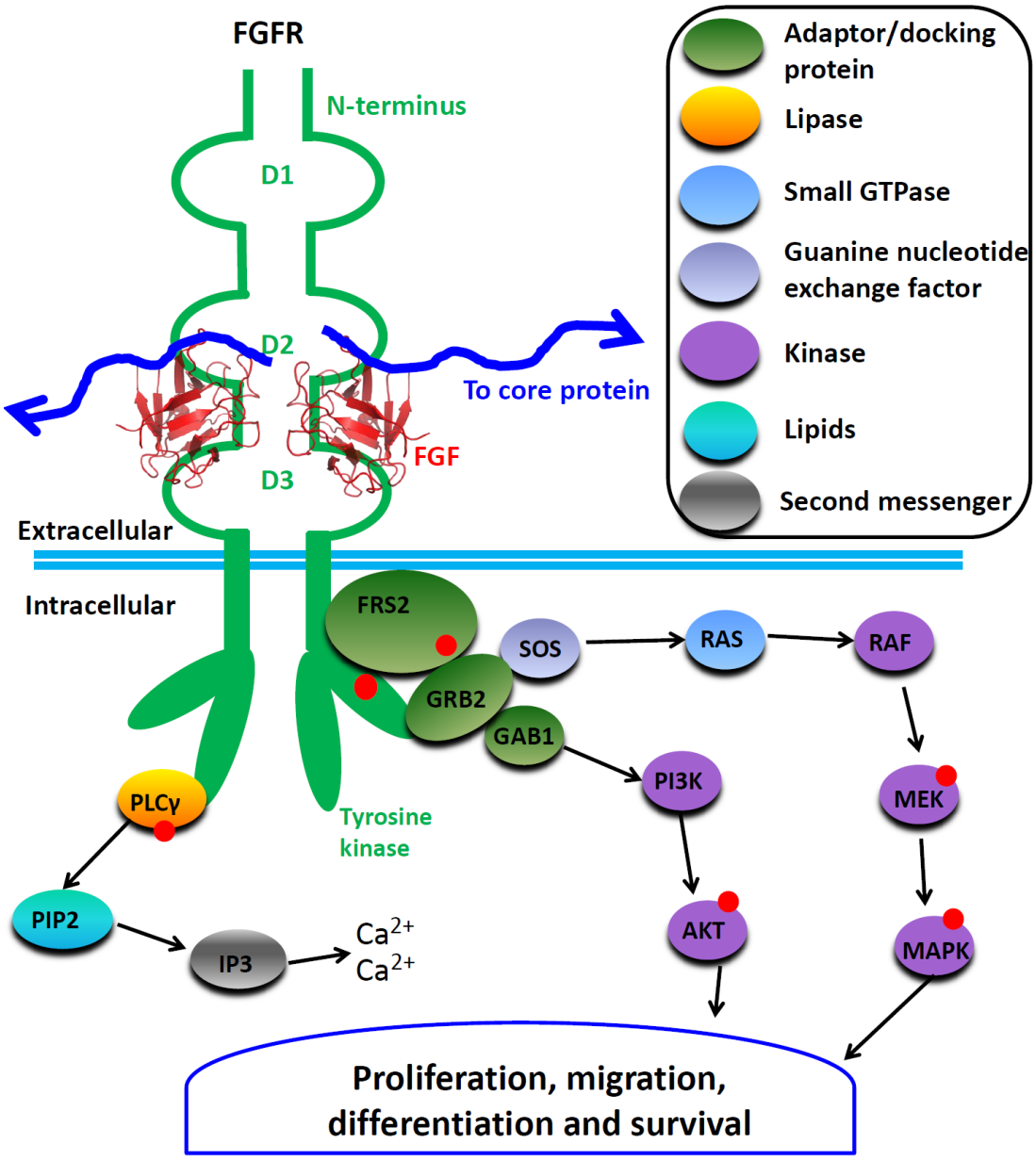


Figure 1.10 FGFR signalling. The interaction of FGF with the FGFR and HS co-receptor leads to the dimerization of the FGFR and the subsequent transphosphorylation of tyrosine residues on the intracellular tyrosine kinase domain. PLCγ and GRB2 can dock to the phosphorylated tyrosines, while FRS2 docks to the juxta-membrane domain. The bound GRB2 activates the RAS-RAF-MAPK and the PI3K-AKT signalling pathways, while phosphorylation of PLCγ induces the phospholipase-IP3 pathway. The signalling pathways regulate many cell activities, e.g., cell fate, cell proliferation, migration and differentiation. FGF, fibroblast growth factor; FGFR, FGF receptor; D1-D3, FGFR immunoglobulin domains 1-3; 'To core protein', indicates that the HS chain (blue line) will be linked to a proteoglycan core

protein; FRS2, FGF receptor substrate 2; GRB2, growth factor receptor-bound protein 2; GAB1, GRB2-associated protein 1; PLCγ, phospholipase Cγ; AKT, Protein kinase B; SOS, son of *sevenless*; RAF: raf-leukemia viral oncogene homologue 1; RAS, rat sarcoma; IP3, inositol trisphosphate; MEK, mitogen-activated protein kinase kinase; MAPK, mitogen-activated protein kinase; PI3K, phosphoinositide 3-kinase; PIP2, phosphatidylinositol 4, 5-bisphosphate; red dots, phosphorylation of signaling proteins.

1.8 Binding and transport of FGFs in ECM

In development, one important feature of paracrine FGFs mediating cell-cell communication is that they will often have to diffuse from the source to the target cells. In addition, this diffusion can set up a gradient (morphogen gradient, Section 1.2). These gradients are a key part of how the fate of cells is specified and how the cell activities are regulated. In the processes of homeostasis and tissue maintenance, FGFs also play important roles. Again, they signal between tissue compartments and so ensure, for example, repair.

Abundant studies have already suggested that the ECMs in different tissues and tissues at different ages or different status are very diverse (Section 1.3) (Frantz et al., 2010, Davies, 2001, Olczyk et al., 2014, Mouw et al., 2014). These ECMs contain different components or same components with different structures, which regulates the binding and diffusion of FGFs. HS chains are the main FGF binding components and by binding to HS chains many spatial and temporal FGF morphogen gradients are formed in development (Duchesne et al., 2012, Dowd et al., 1999, Makarenkova et al., 2009, Shute et al., 2004). Binding to HS also regulates the biological activities stimulated by FGFs in many situations (Zhu et al., 2010,

Delehedde et al., 2000, Delehedde et al., 2002a, Izvolsky et al., 2003, Patel et al., 2008).

The expression of HS chains produces many different tissue-specific structures, which selectively bind to the secreted FGFs and control their diffusion (Izvolsky et al., 2003, Shute et al., 2004). Originally, it was found that FGF2 was stored in endothelial cell ECM and it could be transferred to its cellular receptors to induce cell signals (Vlodavsky et al., 1987). At the tissue level, specific distributions of many FGFs have been observed (Izvolsky et al., 2003, Harada et al., 1999, Gonzalez et al., 1996). A study with FGF2 revealed that FGF2 in many tissues (skin, heart, lung, kidney and intestine) is distributed to defined cells (Gonzalez et al., 1996). Similarly, FGF10 has been found to preferentially bind to urothelium rather than lamina propria in bladder and ureter (Zhang et al., 2006a). Even on one cell, the binding of FGF2 in the pericellular matrix is not homogeneous (Duchesne et al., 2012, Nieves et al., 2015). These studies suggest FGFs selectively bind to HS structures in the ECM, which is consistent with the *in vitro* studies with heparin and its derivatives (Section 1.8) and that the distribution of these structures is spatially controlled (Xu et al., 2012, Ori et al., 2009).

In embryonic development, it seems that FGFs usually diffuse through ECM, rather become trapped and stored there, as in the example of the development of vertebrate limb, in which FGF10 is expressed mesenchymally to regulate the development of the overlying apical ectodermal ridge (Zeller et al., 2009). So, how do the paracrine FGFs move to their target in the ECM, after they are secreted? HS in the ECM was found to control the diffusion of FGF2, which indicates that HS proteoglycans in the ECM might regulate morphogen gradient formation (Dowd et al., 1999). FGF10

bound to heparin acrylic beads could hardly disassociate from heparin and diffuse into either a collagen gel or Matrigel, because of its high affinity for heparin, whereas a mutant FGF10 (R178V), which bound more weakly, diffused out of the heparin bead (Makarenkova et al., 2009). FGF9 diffused more slowly in its targeting ECM than its mutant FGF9Eks which is unable to homodimerise to efficiently interact with HS (Kalinina et al., 2009). These data suggest that HS should selectively bind to FGFs if the FGF is to be allowed to diffuse, or if it is to be stored. The study of the movement of gold nanoparticle-FGF2 in the pericellular matrix of rat mammary (Rama) 27 fibroblasts (cells described in Section 2.6) suggests most FGF2 diffuses slowly by reversible binding to HS chains (Duchesne et al., 2012). Since the binding sites for FGF2 in the HS of pericellular matrix were not evenly distributed, this is likely to contribute to the regulation of the diffusion of FGF2 (Duchesne et al., 2012). However, it is still unknown whether these properties of the binding and transport of FGF2 were a general phenomenon or only true for FGF2. Heparanase, a beta glucuronidase, which cleaves HS in NA and NAS domains, was also identified as an accelerator for the diffusion of FGFs by releasing the S domains to which FGFs will generally be bound (Shute et al., 2004).

1.9 Aims of this project

The aim of the thesis was to establish if HS may generally control the binding and diffusion of FGFs and whether the molecular basis of the interactions of FGFs with the polysaccharide affects FGF diffusion, since this would enable local control of the diffusion of different FGFs by HS. In cases where this has been measured, the paracrine FGFs recognise different, though often overlapping, structures in HS and it

seems that the selectivity of FGFs for HS may follow their phylogenetic relationships (Xu et al., 2012, Xu et al., 2013).

Consequently, FGFs from different subfamilies were used in this thesis, to provide a group of proteins with defined HS binding characteristics, linked to their functional and evolutionary relationships. The first step was to produce FGFs suitable for measuring FGF diffusion. In some cases the specificity of the interactions of the chosen FGFs with HS had not been characterised, so this is determined. A heparin library (comprising different chain lengths and different sulfation modifications) was used as a simple *in vitro* tool to determine the specificity of the interactions of these FGFs with heparin/HS. Fixed Rama 27 fibroblasts were used as the cellular model to study the binding and diffusion of multiple FGFs in the pericellular matrix using imaging technology. Three sub-aims were established and the results are described in Chapters 3 to 5:

i) Producing soluble and biologically active FGFs.

ii) Measuring the heparin/HS binding specificity of FGFs.

iii) Determining the binding and diffusion of FGFs in fixed Rama 27 fibroblast pericellular matrix by using imaging technology.

Chapter 2 Materials and methods

2.1 DNA cloning

2.11 Preparation of competent cells (DH5α, Top 10, SoluBL21 and BL21 plysS)

The desired bacterial strain was grown in 8 mL LB broth (MERCK, Hertfordshire, UK) in a 37ºC shaker overnight. Then, 1 mL bacterial culture was transferred into 100 mL LB broth and the culture was incubated at 37ºC in a shaker for 2~3 hours until OD_{600} was 0.4~0.6. The bacteria were harvested by centrifugation at 4000 g for 10 minutes. The supernatant was discarded and the pellet was resuspended in 20 mL transformation buffer I (30 mM potassium acetate (KOAc), 100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$ and 15% (v/v) glycerol, pH 5.8 adjusted with 0.2 M acetic acid) and incubated on ice for 5 minutes. The bacteria in transformation buffer I were collected by centrifugation and the supernatant was discarded. The cell pellet was gently resuspended with 2 mL transformation buffer II (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl and 15% (v/v) glycerol, pH 6.5 adjusted with 1 M KOH). The cells in transformation buffer II were split into Eppendorf tubes and stored at -80ºC.

2.12 Transformation

Competent cells (*E. coli* DH5α, Top10, SoluBL21, BL21 (DE3) plysS) and plasmids were thawed on ice for 10 minutes. Plasmids (1 µL, 50~300 ng/µL) and 70 µL competent cells were added to a sterile 1.5 mL Eppendorf tube, and the mixture was placed on ice for 30 minutes. Then the tube was heated at 42ºC for 45 seconds, and immediately placed on ice for 2 minutes. LB broth (1mL) culture medium was then

added and the bacterial culture was incubated in a shaker (37°C, 250 rpm). After 60 minutes incubation, the cells were centrifuged and the supernatant was discarded. The cell pellet was resuspended in 100 µL LB culture. Twenty µL cell culture was spread onto a solid LB agar (MERCK) medium plate containing an appropriate antibiotic (50 µg/mL ampicillin sodium salt or 100 µg/mL kanamycin sulfate), and bacteria were incubated in a 37°C overnight. Finally, the plate was usually stored at 4°C until required.

2.1.3 Plasmid amplification in DH5α cells and extraction from the cells

The desired plasmid was transformed into DH5α cells, as described in Section 2.1.2. A single colony containing the plasmid was picked from the agar plate and inoculated into 6~8 mL LB broth medium and the culture was incubated in a 37°C shaker overnight. The bacteria were collected by centrifugation at 4000 g room temperature for 10 minutes and the supernatant was discarded. The plasmid was extracted from the bacteria pellet with QIAprep Spin Miniprep Kit (Qiagen, Manchester, UK) and stored at -20°C.

2.1.4 DNA fragment amplification by Polymerase chain reaction (PCR)

The PCR components, including polymerase (Merck), $MgSO_4$, dNTP, primers, templates, buffer and water, were added to a PCR tube (STARLAB (UK), Ltd, Milton Keynes, UK) by following the protocol in Table 2.1 and mixed. Then, the samples were put into a PCR cycler (TC-PLUS, Bibby Scientific, Staffordshire, UK). The DNA fragment was amplified for 30 cycles with the conditions in Table 2.2. The PCR product was analysed by DNA agarose gel electrophoresis (Section 2.1.5) and stored at -20°C.

Table 2.1 PCR reaction (50 µL) components and volume:

PCR Reaction Components	**Volume/** µL
10 X buffer for KOD Hot Start DNA polymerase	5
25 mM $MgSO_4$	3
dNTP (2 mM each)	5
Sense Primer (30 mM)	2
Anti-sense primer (30 mM)	2
Template DNA (200 ng/µL)	1
KOD Hot Start DNA Polymerase (1 U/µL)	1
PCR Grade water	31
Total reaction volume	50

Table 2.2 PCR cycling conditions:

Cycling conditions	**Temp./** ^oC	**Time**	**Circles**
Stage 1 Activation	98	5 min	1
Stage 2.1 Denature	95	15 s	
Stage 2.2 Annealing	Tm - 5	15 s	
Stage 2.3 Extension	70	40 s	30
Stage 3 Complete Extension	70	10 min	1
Stage 4 Hold	4	+∞	1

2.1.5 DNA agarose gel

TAE (Tris-acetate-EDTA) buffer (10 X, 1 L): 20 mL EDTA (0.5 M, pH 8.0), 48.4 g Tris-Cl, 11.42 mL acetic acid and H_2O.

TAE buffer (1 X) was made by diluting 100 mL 10 X TAE buffer into 900 mL dH_2O. Agarose gels (1.2% (w/v)) were made by dissolving 0.36 g agarose into 30 mL 1 X TAE buffer and heating in a microwave oven. SYBR® Safe DNA Gel Stain (2 µL, Life technologies, Paisley, UK) was added to the gel liquid when its temperature was about 40°C to 50°C. Then the warm gel liquid was added into a gel making kit and allowed to set.

DNA solutions were mixed with 6 X loading dye (Thermo Scientific, Leicestershire, UK) and loaded into the wells and 1 kb DNA ladder (Thermo Scientific) was used to identify the size of the DNA fragments. Agarose gels were run at 100 V in 1 X TAE buffer for 30 minutes.

2.1.6 Restriction enzyme digestion

The plasmid (vector) and PCR fragment (insert) were digested with two desired enzymes (NEB, Hitchin, UK) to open the plasmid and expose the restriction sites, such as NcoI and BamHI. The digestion reaction was set up, as described in Table 2.3 and the buffer was chosen according to that recommended by the manufacturer of the enzymes. The reaction proceeded at 37°C for 2 hours.

Table 2.3 Restriction enzyme digestion setup:

Components	**Volume/** µL
DNA (plasmid or PCR product)	5
Enzyme 1	1
Enzyme 2	1
Buffer (10 X)	2
H_2O	11

2.1.7 DNA extraction from DNA gel

The target DNA fragment was excised from the agarose gel with a scalpel, and the gel slice was weighed in a 1.5 mL Eppendorf tube. Then, the DNA fragment was extracted from the gel by using QIAquick Gel Extraction Kit (Qiagen).

2.1.8 T4 Ligation

The ligation components (Table 2.4) were mixed and the ligation reaction was carried out at room temperature overnight. The ligation product was transformed into Top 10 competent cells to amplify the ligated plasmid.

Table 2.4 T4 ligation setup

Ligation components	**Volume/** μL
Insert	x
Vector	y (50 ng)
T4 ligase Buffer (10 X)	2
T4 ligase	2
H_2O	(20-x-y)

Note: The volumes of insert (x μL) and vector (y μL) are based on different molar ratios (4:1 and 6:1).

2.1.9 In-Fusion Ligation

In-Fusion cloning is a ligation method, which is based on the overlap of base pairs at the ligation sites rather than the cleavage sites. The insert was prepared by PCR, and the vector was opened by enzyme digestion. The PCR product was designed to

possess around 15 bp complementary ends, which overlapped the respective ends of the opened vector. The in-fusion ligation (In-Fusion HD Cloning Kit, Clontech, Takara Bio Europe SAS, Saint-Germain-en-Laye, France) was prepared according to Table 2.5. The reaction was incubated for 15 minutes at 50°C, and then placed on ice for transformation.

Table 2.5 In-Fusion Ligation setup:

In-Fusion Ligation components	**Volume/** µL
5X In-Fusion HD Enzyme Premix	2
Linearized Vector	x (50 ng)
Purified PCR Fragment	y (100 ng)
H_2O	z (z=10-2-x-y)
Total volume	10

Note: The values of x and y are based on the concentration of the DNA fragment.

2.1.10 Colony screening for ligation with PCR and DNA sequencing

The colonies growing on the ligation plate were labelled and used as template in the PCR sample preparation. The PCR products of the predicted size were identified by DNA agarose gel electrophoresis and the corresponding colonies were transformed into 8 mL LB broth medium for plasmid extraction (Section 2.1.3). The extracted plasmid was sequenced across the insert in both directions by GATC Biotech (GATC Biotech Ltd., London, UK).

2.2 Protein expression and purification

The proteins were expressed as described in 'HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors' in Chapter 3 and the collected bacterial pellet was stored at -80ºC before protein extraction (Sections 3.2 and 3.3).

2.3 SDS-PAGE

2.3.1 Sample preparation

Samples were prepared by mixing 30 µL protein solution with 10 µL sample buffer (4 X: 4% (w/v) Sodium dodecyl sulfate (SDS), 40% (v/v) glycerol, 24% (v/v) 1 M Tris-Cl (pH 6.8), 5 % (v/v) β-mercaptoethanol, 0.04% (w/v) bromophenol blue and 3.1 mL water). The mixture was heated at 98ºC for 5 minutes, and centrifuged for 2 minutes.

2.3.2 Gel preparation

The resolving gel for 12% (w/v) SDS-PAGE was prepared according to Table 2.6. The mixture was transferred into a 0.75 mm thickness glass plate assembly, and the resolving gel was covered with 0.5 mL water. When the gel had polymerized, the layer of water was discarded. The stacking gel (Table 2.7) was poured into the glass plate assembly and the sample well comb was inserted into the stacking gel. The gels were ready to use when they set.

Table 2.6 Resolving gel 12% (10 mL):

Components	Volume
Acrylamide/bis-acrylamide stock (30%, w/v)	4.0 mL
Tris-Cl (3 M, pH 8.8)	2.5 mL
Water	3.5 mL
10% (w/v) SDS	100 μL
TEMED	10 μL
Ammonium persulfate (50 mg/ml, fresh)	100 μL

Table 2.7 Stacking gel (10 mL):

Components	Volume
Acrylamide/bis-acrylamide stock (30%, w/v)	1.3 mL
Tris-Cl (1.25 M, pH 6.8)	1.0 mL
Water	7.7 mL
10% (w/v) SDS	100 μL
TEMED	20 μL
Ammonium persulfate (50 mg/ml, fresh)	100 μL

2.3.3 Loading samples and running gel

After the comb was removed the gel assembly was placed in the electrophoresis apparatus. Running buffer (1 X: 14.4 g glycine, 6.04 g Tris-Cl and 1 g SDS for 1 L buffer) was added into the tank to cover the gel. Samples (Section 2.3.1) were added into the wells, and the gel was run at 30 mA per gel (maximum voltage: 200 V).

When the dye front reached the bottom of the gel, the gel was transferred into a staining box.

2.3.4 Coomassie staining

The gel was stained in staining buffer (300 mL methanol, 100 mL acetic acid, 550 mL water, 50 mL glycerol and 1 g Coomassie blue) on a shaker (30 rpm) for 30 minutes. Then, the staining buffer was discarded and the gel was destained with destaining buffer (300 mL methanol, 100 mL acetic acid and 600 mL water) on a shaker until the background was clear. The destained gel was imaged with a scanner.

2.4 Cell disruption and protein purification

The bacterial cells containing the expressed protein were disrupted and the soluble fraction was obtained and the expressed FGFs were purified as described, in 'HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors', in Chapter 3.

2.5 Differential Scanning Fluorimetry (DSF)

The experimental procedure, data acquisition and data analysis are described in Chapter 4 (Section 4.2).

2.6 Tissue culture

2.6.1 Routine culture

Cell culture medium for Rama 27 fibroblasts (Rudland et al., 1984) and HaCaT keratinocytes (Boukamp et al., 1988) were prepared as described in 'HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors' in Chapter 3. The Rama 27 cells were isolated form the fast sticking fraction obtained from rat mammary gland and were triply single cell cloned. They were designated as fibroblasts based on their ability to be differentiated to a pre-adipocyte phenotype (Rudland et al., 1984).

The cells were cultured in a 37ºC incubator with 10% (v/v) CO_2 until the 10 cm culture plate (Corning, Nottingham, UK) was ~90% confluent. The plate was washed with PBS (2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 and 0.15 M NaCl, pH 7.4) twice and then incubated with 1 mL versene (Life technologies) containing 2.5% (v/v) trypsin stock (2.5% (w/v), Life technologies) in the incubator to release cells. The released cells were resuspended with 8 mL culture medium and split into 8 dishes (Rama 27) or 4 dishes (HaCaT). Another 8 mL culture medium was added to each dish and the dishes were returned to the incubator.

2.6.2 Counting cells

Cells released by trypsinisation (Section 2.6.1) (0.5 mL) in cell culture medium (total volume: 10 mL) were added to a Coulter counting cup containing 19.5 mL Isoton II (Beckman Coulter, High Wycombe, UK). The cell number in 0.5 mL solution in the cup was determined in a Z1 coulter particle counter (Beckman Coulter). The total

number of cells in the dish was calculated by accounting for the dilutions, so by multiplying the number by 800.

2.6.3 Freezing cells

The cells were released from the cell dish by trypsin digestion (Section 2.6.1) and 9 mL cell culture medium was added. The cells were counted, placed in a 20 mL universal tube and collected by centrifugation at 800 rpm for 5 minutes. The supernatant was discarded and the cell pellet was resuspended in freezing medium (14.5 mL Dulbecco's modified Eagle's medium (DMEM, Life technologies), 1.5 mL dimethyl sulfoxide and 4 mL fetal calf serum (FCS, Labtech International Ltd, East Sussex, UK)) to obtain the desired cell density (1~1.5 x 10^6 cells/mL). The cells (1 mL) in freezing medium were split into 1.5 mL cryotubes and transferred into a -80^{o}C freezer. The frozen cells were stored in a -140^{o}C freezer.

2.6.4 Thawing cells

The cells taken out from the freezer were thawed in a 37^{o}C water bath. Then, the cells were transferred into a universal tube and 20 mL thawing medium (16 mL DMEM medium and 4 mL FCS) was slowly added. The cells in the thawing medium were collected by centrifugation at 800 rpm for 5 minutes and 10 mL cell culture medium was added to resuspend the cells, which were then placed into a 10 cm dish and transferred to the 37^{o}C incubator.

2.6.5 Cell growth assay and p44/42^{MAPK} signalling assay

The cell growth assay and Western blotting for p44/42^{MAPK} were performed as described, in 'HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors', in Chapter 3.

2.7 Fluorescence imaging

The sample preparation, imaging and data analysis were performed, as described in Sections 5.2, 5.3 and 5.4 in Chapter 5.

Chapter 3 Production of FGFs in *E. coli*

3.1 Introduction

An important aim of this thesis was to produce a range of FGFs from different subfamilies so that their interactions with GAGs could be investigated (Sections 1.5.4 and 1.9). A number of FGFs had previously been expressed in *E. coli* and purified to explore their structures and interactions with their receptors (FGFR) and their co-receptors (Heparin/HS) (Zhang et al., 2006b, Xu et al., 2012, Xu et al., 2013). However, the existing set of FGFs was limited and only one FGF was available from most subfamilies and none from the FGF4 subfamily was available (Xu et al., 2012). To test the hypothesis that the specificity of interactions of FGFs with GAGs was under the same natural selection pressure that drove the diversification of the FGFs, it would require a more comprehensive set of proteins. The functional implications of this hypothesis would also necessitate a larger set of FGFs.

One reason for the previous expression of a limited set of FGFs was that many were insoluble. Consequently, a major focus was to develop the means of producing soluble proteins, which included testing a variety of bacterial hosts and the SUMO solubilisation tag. Alongside, a method to label FGFs with high structural specificity was required for fluorescence microscopy. HaloTag, a mutated dehalogenase (Los et al., 2008), was tested in this respect, since it potentially would provide great versatility in terms of labelling; the N-terminus of FGFs was deemed to be a suitable position for the fusion protein, since FGF2 had successfully been labelled at that

position with gold nanoparticles (Duchesne et al., 2012). In this chapter the production of recombinant FGFs with and without an N-terminal HaloTag is described. The activity of these FGFs was also measured, since this depends on both binding to HS and FGFR and so would determine if the FGFs were properly folded. Much of this work is published in the following paper, with full author contributions included.

3.2 Paper

Sun, C., Li, Y., Taylor, S. E., Mao, X., Wilkinson, M. C., Fernig, D. G. (2015). HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors. *PeerJ*. **3**, e1060.

Contributions of the authors:

Changye Sun: Cloning, design of cloning strategy, expression and purification of His-FGF1, His-FGF10, His-FGF6, His-FGF8, His-FGF20, His-FGF22, Halo-FGF6, Halo-FGF8, Halo-FGF20 and Halo-FGF22, measurement of protein yields, of stimulation of phosphorylation of p42/44^{MAPK} and of cell growth by FGFs, planed and wrote the paper. **With Sarah E. Taylor** cloning, expression and purification of Halo-FGF2; **With Xianqing Mao** cloning, expression and purification of Halo-FGF3, Halo-FGF7 and Halo-FGF10; **With Yong Li cloning**, expression and purification of Halo-FGF16 and Halo-FGF17; expression and purification of FGF2, His-FGF3, His-FGF7 (Plasmids from Ruoyan Xu, University of Liverpool); **Mark C. Wilkinson:** Provided guidance on protein purification and edited the paper; **David G. Fernig:** conceived the study and edited the paper.

PeerJ

HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors

Changye Sun[1,3], Yong Li[1,3], Sarah E. Taylor[1], Xianqing Mao[2], Mark C. Wilkinson[1] and David G. Fernig[1]

[1] Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool, UK
[2] Department of Oncology, Laboratory of Cellular and Molecular Oncology, Luxembourg Institute of Health, Luxembourg
[3] These authors contributed equally to this work.

ABSTRACT

The production of recombinant proteins such as the fibroblast growth factors (FGFs) is the key to establishing their function in cell communication. The production of recombinant FGFs in *E. coli* is limited, however, due to expression and solubility problems. HaloTag has been used as a fusion protein to introduce a genetically-encoded means for chemical conjugation of probes. We have expressed 11 FGF proteins with an N-terminal HaloTag, followed by a tobacco etch virus (TEV) protease cleavage site to allow release of the FGF protein. These were purified by heparin-affinity chromatography, and in some instances by further ion-exchange chromatography. It was found that HaloTag did not adversely affect the expression of FGF1 and FGF10, both of which expressed well as soluble proteins. The N-terminal HaloTag fusion was found to enhance the expression and yield of FGF2, FGF3 and FGF7. Moreover, whereas FGF6, FGF8, FGF16, FGF17, FGF20 and FGF22 were only expressed as insoluble proteins, their N-terminal HaloTag fusion counterparts (Halo-FGFs) were soluble, and could be successfully purified. However, cleavage of Halo-FGF6, -FGF8 and -FGF22 with TEV resulted in aggregation of the FGF protein. Measurement of phosphorylation of p42/44 mitogen-activated protein kinase and of cell growth demonstrated that the HaloTag fusion proteins were biologically active. Thus, HaloTag provides a means to enhance the expression of soluble recombinant proteins, in addition to providing a chemical genetics route for covalent tagging of proteins.

Subjects Biochemistry
Keywords Fibroblast growth factor, Recombinant protein expression, HaloTag, Fusion protein

Submitted 23 December 2014
Accepted 8 June 2015
Published 25 June 2015

Corresponding author
David G. Fernig, dgfernig@liv.ac.uk

Academic editor
Sandhya Visweswariah

Additional Information and Declarations can be found on page 21

DOI **10.7717/peerj.1060**

Copyright
2015 Sun et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

INTRODUCTION

Of the 18 receptor-binding fibroblast growth factors (FGF), 15 also bind a heparan sulfate co-receptor and are classed as growth factors and morphogens. These are grouped into 5 subfamilies based on their protein sequence similarity (*Itoh, 2007*; *Ornitz, 2000*), and they regulate a myriad of processes in development, homeostasis and in some diseases (*Beenken & Mohammadi, 2009*; *Turner & Grose, 2010*). Recombinant FGFs provide a key tool to study their structure–function relationships, and labelling FGFs for microscopy

How to cite this article Sun et al. (2015), HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors. **PeerJ 3:e1060; DOI 10.7717/peerj.1060**

has been important in probing the mechanisms of, for example, their transport (*Duchesne et al., 2012*; *Lin, 2004*; *Yu et al., 2009*). Chemical labelling has disadvantages compared to genetically encoded labelling, since with the latter it is easier to predict the structural and hence functional consequences of labelling, which can be achieved both *in vitro* and *in vivo*. While fluorescent proteins remain a mainstay of genetic labelling, they have limitations. These have been overcome, for example, by non-covalent tagging of proteins on hexahistidine sequences with Tris-Ni^{2+} nitriloacetic acid (*Huang et al., 2009*; *Lata et al., 2005*; *Tinazli et al., 2005*), which has allowed diverse labelling strategies, ranging from fluorescent dyes (*Uchinomiya et al., 2009*) and quantum dots (*Roullier et al., 2009*; *Susumu et al., 2010*) to gold nanoparticles (*Duchesne et al., 2008*). However, non-covalent coupling is reversible and exchange may occur in this instance with histidine-rich patches on endogenous proteins.

HaloTag is a mutant of a bacterial haloalkane dehalogenase, which reacts with chloroalkane ligands to form a covalent bond that represents the covalent intermediate of the enzyme's normal catalytic cycle (*Los et al., 2008*). Fluorescent dyes (*Los et al., 2008*) and quantum dots (*Zhang et al., 2006b*) carrying a chloroalkane group have been used to label HaloTag fusion proteins for fluorescence imaging. This approach is particularly versatile, since it combines the power of a genetically encoded tag (the HaloTag protein) with covalent labelling.

Consequently, we set out to produce N-terminal HaloTag fusions of different FGFs. In the course of this work, we observed that the N-terminal HaloTag fusion had a substantial effect on the expression of the more recalcitrant FGFs, consistent with the observation that HaloTag is a potential solubilisation tag for recombinant proteins (*Ohana et al., 2009*). Thus, whereas expression of FGF1 and FGF10 was somewhat reduced and that of FGF2 increased, expression of FGF7, which can be toxic (*Ron et al., 1993*), was no longer so, while expression of soluble FGF3, FGF6, FGF7, FGF8, FGF16, FGF17, FGF20 and FGF22 was markedly enhanced. This is in contrast to previous reports where FGFs such as FGF6 (*Pizette et al., 1991*), FGF8 (*Loo & Salmivirta, 2002*; *Macarthur et al., 1995*; *Vogel, Rodriguez & IzpisuaBelmonte, 1996*), FGF16 (*Danilenko et al., 1999*) and FGF20 (*Jeffers et al., 2002*; *Kalinina et al., 2009*) have been found to be mainly expressed in inclusion bodies, even as truncated proteins, and so require refolding. Thus, HaloTag provides not just a means to label proteins covalently and specifically, but is also a useful solublisation partner for the production of recombinant proteins.

MATERIALS AND METHODS

Materials

pET-14b vectors containing cDNAs encoding FGF1 and FGF2 and pET-M11 vector containing FGF7 cDNA were as described (*Xu et al., 2012*); cDNAs encoding FGF3, FGF10, FGF16, FGF17 and FGF20 were purchased from Eurofins Genomics (Ebersberg, Germany); cDNA encoding FGF6, FGF8 and FGF22 were purchased from Life Technologies (Paisley, UK); cDNAs encoding HaloTag was acquired from Kazusa DNA Research Institute (Kisarazu, Japan); Primers for PCR were from Life Technologies (Paisley, UK). All

Table 1 Peptide sequences of FGFs, the N-terminal HisTag constructs and the N-terminal HaloTag constructs. FGF names, sequences and amino acid numbering are according to the UniProt entry. FGF1 is an N-terminal truncated protein (*Ke et al., 1990*). FGF2 does not possess a secretory signal sequence, whereas there is no signal peptide recognised in Uniprot for FGF16 and FGF20; consequently full length protein sequence was expressed. For all other FGFs, the protein expressed was without the Uniprot determined secretory signal sequence. FGFx refers any one of the FGFs. TEV cleavage sites are in red.

Name	UniProt accession number	Residues in mature protein
FGF1	P05230	16–155
FGF2	P09038-2	1–155
FGF3	P11487	18–239
FGF6	P10767	38–208
FGF7	P21781	32–194
FGF8b	P55075-3	23–215
FGF10	O15520	38–208
FGF16	O43320	1–207
FGF17	O60258-1	23–216
FGF20	Q9NP95	1–211
FGF22	Q9HCT0	23–170
HisTag terminus (pET-M11)		MKHHHHHHPMSDYDIPTTENLYFQGA-[FGFx]
HaloTag and TEV site to conjoin with FGF sequence		MPEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYV WRNIIPHVAPTHRCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEAL GLEEVVLVIHDWGSALGFHWAKRNPERVKGIAFMEFIRPIPTWDEWPE FARETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVVRPLTEVEMDHYREP FLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKLLFWG TPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLS TLEISGEPTTEDLYFQS-[FGFx]

of the protein sequences corresponding to the above cDNAs are listed in Table 1. Enzymes for cloning were from: NcoI, BamHI and T4 ligase (NEB, Hitchin, UK); KOD Hot Start DNA polymerase (Merck, Hertfordshire, UK); In-Fusion® HD Cloning Kit (Clontech, Takara Bio Europe SAS, Saint-Germain-en-Laye, France). Bacterial cells: DH5α, BL21 (DE3) pLysS and SoluBL21 were a gift from Olga Mayans, University of Liverpool. The sources of other materials were as follows: LB broth and LB agar (Merck, Hertfordshire, Germany); Soniprep 150 Plus (MSE, London, UK); Affi-Gel® Heparin Gel (Bio-Rad, Hertfordshire, UK), CM Sepharose Fast Flow, DEAE Sepharose Fast Flow, HiTrap Q HP column; empty disposable PD-10 Columns; ÄKTApurifier 100 plus (GE Healthcare, Buckinghamshire, UK).

For cell culture the following materials were used: Dulbecco's modified Eagle's medium (DMEM, Life Technologies), fetal calf serum (FCS, Labtech International Ltd, East Sussex, UK), 7.5% (w/v) sodium bicarbonate (Invitrogen, Paisley, UK), 200 mM L-glutamine (Gibco), 5 µg/mL insulin (Sigma-Aldrich, Dorset, UK), 5 µg/mL hydrocortisone (Sigma-Aldrich), bovine serum albumin (BSA; A7030, Sigma-Aldrich) for cell culture, cell culture dishes (Corning, Nottingham, UK). For SDS-PAGE and Western blotting: dried skimmed milk (Marvel, Spalding, UK), BSA (Fisher Scientific, Lougborough, UK), protease inhibitor (Roche, Burgess Hill, UK), phospho-p44/42 MAPK (T202/Y204) antibody (Cell Signalling, NEB, Hitchin, UK), monoclonal β-actin antibody

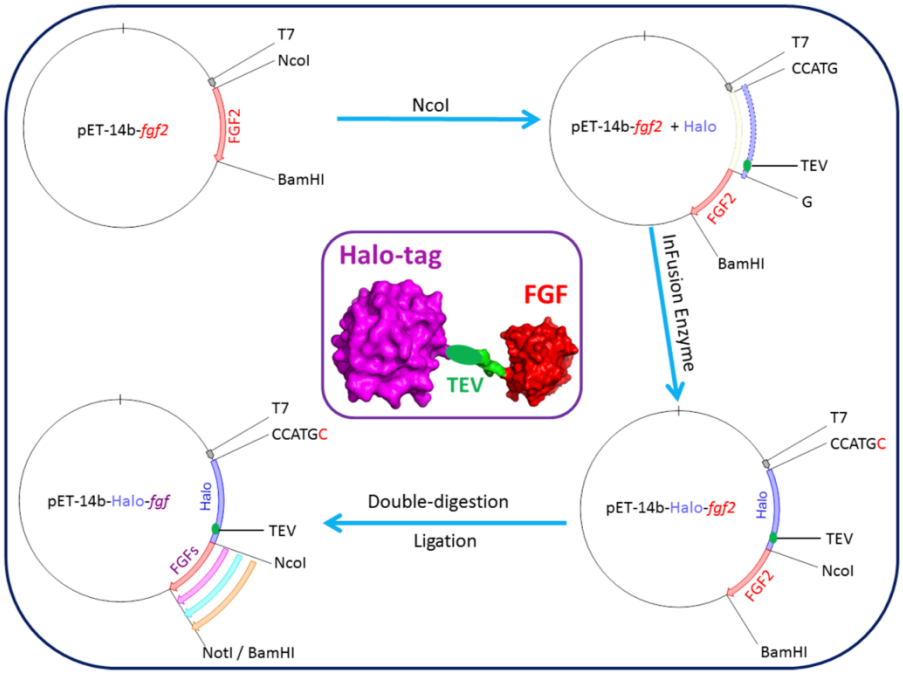


Figure 1 Cloning strategy for plasmids encoding Halo-FGFs. DNA encoding HaloTag was inserted 5′ of the FGF2 coding sequence with the In-Fusion HD enzyme. Subsequently, a NotI cleavage site was added 5′ to the BamHI site and other FGFs were exchanged into the plasmid using the digestion-ligation cloning method. A cartoon structure of Halo-FGF is presented in the middle of this figure.

(Sigma-Aldrich, Dorset, UK), anti-mouse IgG, horseradish peroxidase-linked antibody (Cell Signalling, NEB), polyvinylidene fluoride (PVDF) transfer membrane (Millipore UK, Hertfordshire, UK), enhanced chemiluminescence (ECL) Western blotting reagents (GE Healthcare, Little Chalfont, UK), Hyperfilm (GE Healthcare, Little Chalfont, UK).

DNA cloning of hexahistidine tagged FGFs (His-FGFs) and HaloTag tagged FGFs (Halo-FGFs)

DNA encoding FGF1, FGF3, FGF6, FGF8, FGF10, FGF16, FGF17, FGF20 and FGF22 was cloned into pET-M11 such that the resulting protein would have a N-terminal 6xhis tag followed by a tobacco etch virus (TEV) cleavage site (ENLYFQ). FGF2 and FGF7 DNA sequences were previously cloned into pET-14b and pET-M11, respectively (*Xu et al., 2012*).

A plasmid encoding Halo-FGF2 was produced by adding a HaloTag encoding DNA sequence in-frame 5′ to a DNA sequence encoding full-length FGF2. This construct was then used to produce the other DNAs encoding Halo-FGFs (Fig. 1). The plasmid pET-14b-*fgf2* contains NcoI and BamHI cleavage sites 5′ and 3′ of *fgf2*, respectively. This vector was linearized by digestion with NcoI. The DNA encoding HaloTag (Fig. 1: blue insert) was amplified by PCR using the Halo-FGF2-Forward, AAGGAGATATA CCATGCCAGAAATCGGTACTG, and Halo-FGF2-Reverse,

Table 2 Concentrations of NaCl in 50 mM Tris-Cl buffer (pH 7.4) used for heparin affinity chromatography of FGFs. [NaCl] for lysate is the concentration of NaCl in the sample applied to the column.

Name	[NaCl] for lysate (M)	[NaCl] for wash (M)	[NaCl] for elution (M)
FGF1	0.6	0.6	1.0
FGF2	0.6	0.6	1.5
FGF3	0.3	0.6	1.0
FGF6	0.3	0.4	1.0
FGF7	0.3	0.3	1.0
FGF8	0.6	0.6	1.5
FGF10	0.6	0.6	1.0
FGF16	0.3	0.4	1.0
FGF17	0.6	0.6	1.0
FGF20	0.3	0.4	1.0
FGF22	0.6	0.8	1.5

TCCCGGCTGCCATGGAGCTCTGAAAGTACAGATC, primers (NcoI cleavage site underlined), and inserted into the linearized vector using In-Fusion enzyme. A TEV cleavage site (Fig. 1: green ellipsoid) was also included at the C-terminus of HaloTag to allow release of the FGF. A NotI cleavage site was also inserted 5′ of the BamHI to provide an additional 3′ cleavage sites for cloning. The other cDNAs (FGF1, FGF3, FGF6, FGF7, FGF8, FGF10, FGF16, FGF17, FGF20 and FGF22) were exchanged into the established pET-14b-*Halo-fgf2* plasmid by double-digestion with NcoI and BamHI/NotI enzymes and ligation using T4 ligase (Fig. 1).

Protein expression and purification of His-FGFs and Halo-FGFs

His-FGF7, because it is toxic like native FGF7 (*Ron et al., 1993*), was transformed into BL21 (DE3) pLysS (F– ompT hsdSB(rB–, mB–) gal dcm (DE3) pLysS (CamR)) for subsequent protein expression and purification. FGF2, the other His-FGFs and Halo-FGFs were transformed into SoluBL21(F– *ompT hsd*SB(rB–, mB–) *gal dcm* (DE3)). The bacteria containing FGF encoding plasmids were cultured at 37 °C until the OD600 values were between 0.4 and 0.6, and then protein expression at 16 °C was induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The bacteria were harvested by centrifugation at 4 °C, 14,000 g for 10 min and the pellets frozen at −80 °C.

The bacterial pellets were resuspended with the corresponding 50 mM Tris-Cl lysate buffers (pH 7.4) (Table 2), and the cells were disrupted by 5–6 cycles of sonication (30 s sonication, 60 s pause) on ice. Cell debris and insoluble proteins were removed by centrifugation at 4 °C, 30,000 g for 30 min. Then, the presence of soluble FGFs was tested by analysis of whole cells, the supernatant and pellet by separation of polypeptides on 12% (w/v) SDS-PAGE and coomassie staining.

FGF2 and His-FGF7 were purified as described before (*Xu et al., 2012*). Soluble FGF1, FGF2, FGF3, FGF10, FGF16 and FGF17, including His-FGFs and Halo-FGFs, were loaded onto a 3 mL and the other soluble FGFs were loaded onto an 8 mL column of heparin

agarose. For each FGF, different concentrations of NaCl (in 50 mM Tris-Cl pH 7.4) were used for washing and elution (Table 2) by following the previous measurements on the electrolyte sensitivity of their heparin binding assessed by Western blot (*Asada et al., 2009*). The yields of His-FGFs and Halo-FGFs were quantified by measuring the absorbance at 280 nm and the level of impurities were estimated by analysis of coomassie stained SDS-PAGE gels with ImageJ-Analyze-Gels (*Ferreira & Rasband, 2012*). The soluble His-FGFs eluted from heparin affinity chromatography were further purified by Ni^{2+} affinity chromatography. Due to the negative charge on the surface of HaloTag and positive charge on the surface of FGFs, Halo-FGFs could bind to both cation- and anion-exchange stationary phases. Thus, Halo-FGF1, Halo-FGF2, Halo-FGF3, Halo-FGF7 and Halo-FGF10 were purified by chromatography on a 5 mL HiTrap Q HP column. Samples were applied in 0.15 M NaCl in PB buffer (2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) and eluted with a gradient running to 0.8 M NaCl in the same buffer. Halo-FGF6 and Halo-FGF20 were purified by chromatography on a 3 mL column of CM Sepharose Fast Flow followed by a 3 mL column of DEAE Sepharose Fast Flow. Samples were again applied in 0.15 M NaCl in PB buffer and eluted with 0.4 M NaCl in the same buffer. The purified His-FGFs and Halo-FGFs were analysed by 12% (w/v) SDS-PAGE followed by coomassie staining.

Purification of FGFs by removing HaloTag from Halo-FGFs

To test the accessibility of the TEV cleavage site, some Halo-FGFs, including Halo-FGF2, Halo-FGF17, Halo-FGF6, Halo-FGF8 and Halo-FGF22 eluted with high concentration of NaCl in 50 mM Tris buffer from heparin agarose chromatography and Halo-FGF20 purified with heparin, DEAE and CM chromatography, were incubated with 2.5% (mol/mol) TEV protease at 4 °C overnight. In cases where the digestion products were cloudy, they were clarified by centrifugation for 30 min at 13,000 g, 4 °C. Samples were then analysed on a 12% (w/v) SDS-PAGE. The supernatants of the TEV digestions of Halo-FGF6 and of Halo-FGF20 were applied onto a 2 mL heparin agarose column, and washed as before (Table 2). FGF6 and FGF20 were eluted with PB buffer containing 1 M NaCl or 0.1 M arginine and 1 M NaCl, respectively. After TEV digestion, FGF17 was further purified on a 1 mL HiTrap SP HP cation-exchange column by washing with 0.3 M NaCl and eluting with 1 M NaCl, both in 50 mM Tris-Cl, pH 7.4. All of the fractions from the purification steps were analysed by 12% (w/v) SDS-PAGE.

Cell culture

Rama 27 cells were cultured in DMEM medium containing 10% (v/v) FCS, 4 mM L-glutamine, 0.75% (w/v) sodium bicarbonate, 50 ng/mL insulin and 50 ng/mL hydrocortisone (*Rudland, Twiston Davies & Tsao, 1984*). HaCaT cells were cultured in the same medium, but without insulin and hydrocortisone (*Boukamp et al., 1988*). Cell number was measured with a Z1 coulter particle counter (Beckman Coulter, High Wycombe, UK).

Measurement of p44/42^{MAPK} phosphorylation

Cells were cultured in 3 cm dishes until near confluence. Then, the dishes were washed twice with phosphate-buffered saline (PBS) and 2.5 mL step-down medium (SDM: DMEM with 250 ng BSA, 0.75% (w/v) sodium bicarbonate and 4 mM L-glutamine) was added for 24 h (Rama 27 cells) or 48 h (HaCaT cells). Rama 27 and HaCaT cells were then incubated with different FGFs for 15 min, as described in the figure legends. After the incubation, the cells were washed twice with ice-cold PBS and collected by scraping in 2X SDS-PAGE lysis buffer (4% (w/v) SDS, 20% (v/v) glycerol, 12% (v/v) Tris-Cl (pH 6.8), 2.5% (v/v) β-mercaptoethanol, 0.02% (w/v) bromophenol blue, 1 tablet of protease inhibitor and 6.8 mL distilled water). The cell lysates were heated for 10 min at 98 °C prior to SDS-PAGE.

Western blot

After separation by 10% (w/v) SDS-PAGE, polypeptides were transferred onto a PVDF membrane. The membrane was blocked with 5% (w/v) skimmed milk in 1X TBST (50 mM Tris-Cl, 150 mM NaCl and 0.05% Tween-20 (v/v), pH 7.5) for 2 h. After two washes with TBST, the membrane was incubated with phospho-p44/42^{MAPK} antibody (1:1,000 dilution in TBST) on a shaker overnight at 4 °C. Secondary anti-mouse antibody (1:1,000 dilution) was added to the membrane after three washes with TBST, 5 min each, for 1 h at room temperature. Following three washes with TBST to remove the excess secondary antibody, the membrane was covered with 1 mL ECL solution and signal was detected with Hyperfilm. The same membrane was stripped with 2.5% (w/v) SDS in TBST at 50 °C for 1 h and reblocked as above, before probing with β-actin antibody (1:10,000 dilution). The Western blot band intensity was quantified in the same way as SDS-PAGE bands and the signal intensities of phospho-p44/42^{MAPK} were normalised by dividing by the intensity of the band corresponding to β-actin and then by that of the BSA control samples.

Cell growth assay

Cell growth in Rama 27 fibroblasts was measured as before (*Smith, Winslow & Rudland, 1984*). Rama 27 cells were dispensed into a 24 well cell culture plates at 2,000 cells/well. After 24 h the cells were washed twice with PBS and cultured in SDM, as described for the p44/42^{MAPK} phosphorylation assay for 24 h. The SDM was then replaced and the appropriate proteins added, as described in the figure legend. After 68 h incubation, cells were trypsinised and the number of cells counted.

RESULTS AND DISCUSSION

Expression of soluble FGFs

Based on their relative expression and solubility properties, the FGFs were split into three different groups: FGFs that expressed well as soluble proteins (Group 1: FGF1, FGF2 and FGF10), FGFs that expressed at a low level (Group 2: FGF3 and FGF7), and FGFs that were insoluble when expressed in *E. coli* (Group 3: FGF6, FGF8, FGF16, FGF17, FGF20 and FGF22).

Group 1: soluble FGFs

After induction, bands corresponding to the expected molecular size of His-FGF1, FGF2 and His-FGF10 were apparent in the whole cell lysates (Figs. 2A, 2C and 2E, lane L, green arrow). His-FGF1 and His-FGF10 were expressed at a higher level than FGF2 in *E. coli* SoluBL21. After centrifugation of the cell lysates, bands corresponding to the molecular size of all three FGFs were mainly recovered in the soluble fraction (supernatant), rather than in the insoluble fraction (pellet; Figs. 2A, 2C and 2E, lanes S and P). Chromatography of the supernatants on heparin demonstrated that little of the expressed protein was present in the flow-through fraction (Figs. 2A, 2C and 2E, lane T). Weak bands corresponding to His-FGF1 and His-FGF10, but not FGF2, were observed in the wash fraction (Figs 2A and 2E, lane Wa), which may represent aggregated or less well-folded protein. The majority of the three FGFs was recovered in the high NaCl eluate (Figs. 2A, 2C and 2E, lane Hep), which demonstrated that these soluble FGFs bound heparin strongly. This indicated that they were likely to be properly folded, because the canonical, highest affinity heparin binding site of FGFs depends on the tertiary structure of the proteins (*Xu et al., 2012*).

The bands corresponding to Halo-FGF1, Halo-FGF2 and Halo-FGF10 were clearly observed in the whole cell lysates and these proteins were all highly expressed in SoluBL21 cells (Figs. 2B, 2D and 2F, lane L, red arrow). Similarly to the His-FGF1, FGF2 and His-FGF10, after centrifugation of the whole cell lysates, the bands corresponding to the three Halo-FGFs were observed in the soluble fractions (Figs. 2B, 2D and 2F, lanes S and P). Chromatography of the soluble fractions on heparin indicated that most of Halo-FGF2 and Halo-FGF10 had bound to the column, but there was a substantial amount of Halo-FGF1 in the flow-through (Figs. 2B, 2D and 2F, lane T). This may be due to the capacity of the column for Halo-FGF1 being lower than for His-FGF1. All three Halo-FGFs were eluted from the heparin affinity column at the expected NaCl concentration (Figs. 2B, 2D and 2F, lane Hep).

The yield of Halo-FGF1 and Halo-FGF10 was similar to that of the corresponding his-tagged proteins (Table 3). However, since the Halo-FGF proteins are considerably larger than the corresponding His-tagged FGF1 and FGF10, this represents a decrease in the molar amounts of FGF produced. In contrast, the yield of Halo-FGF2 was 4-fold higher (Table 3), which is only partly accounted for by the increased size of the fusion protein. The low yield of full-length FGF2 has been ascribed to the presence of secondary structure at the 5′ end of the FGF2 mRNA (*Knoerzer et al., 1989*), and the presence of the upstream HaloTag sequence may mitigate this effect.

Group 2: low expression proteins

The expression of His-FGF3 was weak, as was that of His-FGF7 (expressed in BL21 DE3 pLysS) due to its toxicity (*Ron et al., 1993*) (Figs. 3A and 3C, lane L, S and P, green arrow). Heparin chromatography of the supernatants demonstrated that the yields of soluble His-FGF3 and His-FGF7 were quite low (Figs. 3A and 3C, lane Hep; Table 3).

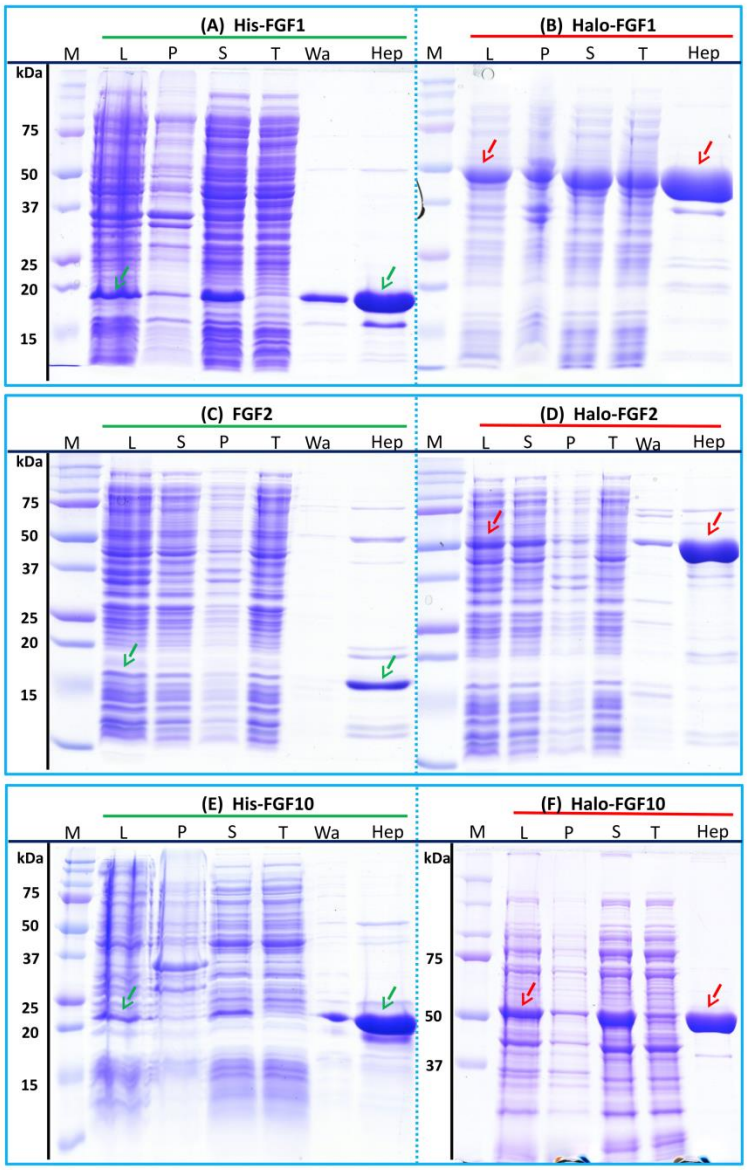


Figure 2 Expression and heparin affinity purification of His-FGF1, FGF2, His-FGF10, Halo-FGF1, Halo-FGF2 and Halo-FGF10. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The supernatant was subjected to heparin-affinity chromatography and samples were then analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant; T, unbound, flow-through fraction from heparin-affinity chromatography; Wa, wash of heparin-affinity column (Table 2); Hep, high NaCl eluate of heparin-affinity column (Table 2). Green arrows: FGF or His-FGF; red arrows: Halo-FGF.

Table 3 Summary of the molecular sizes and yields of His-FGFs and Halo-FGFs. The molecular weight of the proteins was calculated from their amino acid sequence. The concentrations and volumes of His-FGFs and Halo-FGFs recovered from heparin affinity chromatography were measured. The impurities identified by SDS-PAGE were quantified using ImageJ relative to the band corresponding to His-FGF and to Halo-FGF and the amount of protein in the eluate from heparin chromatography adjusted accordingly, to provide an estimate of the yield.

FGFs	Molecular weight (kDa)		Yield (mg/L)	
	HisTag	HaloTag	HisTag	HaloTag
FGF1	19.1	50.9	14	16
FGF2	17.3 No Tag	52.2	2.5 No Tag	11
FGF3	28.2	60.0	0.5	11
FGF6	22.3	54.1	n.d.[a]	27
FGF7	22.2	54.0	0.6	5.6
FGF8	25.7	57.5	n.d.[a]	1.7
FGF10	22.7	54.5	7.7	9.3
FGF16	26.9	58.7	n.d.[a]	1.0
FGF17	25.8	57.6	n.d.[a]	1.5
FGF20	26.9	58.6	n.d.[a]	10
FGF22	20.5	52.3	n.d.[a]	2.0

Notes.

[a] Not detected. Insufficient soluble protein for reliable quantification.

Transformation of SoluBL21with the plasmid encoding Halo-FGF7 yielded the expected number of colonies, indicating that the fusion protein was not toxic. Bands corresponding to the molecular size of Halo-FGF3 and Halo-FGF7 were observed in the cell lysates (Figs. 3B and 3D, lane L, red arrow) and in the soluble fraction obtained after centrifugation, whereas the pellet has relatively weaker bands (Figs. 3B and 3D, lanes P and S), indicating that Halo-FGF3 and Halo-FGF7 were soluble. Heparin chromatography of the soluble factions demonstrated that large amounts of Halo-FGF3 and Halo-FGF7 retained their heparin binding interaction with the polysaccharide (Figs. 3B and 3D, lane Hep).

The yields of Halo-FGF3 and of Halo-FGF7 were 21-fold and 9-fold greater than of the corresponding His-tagged FGF (Table 3). Thus, the presence of the HaloTag N-terminal fusion increased the amounts of FGF3 and FGF7 substantially, even after taking into account the larger size of these fusion proteins (Table 3).

Group 3: insoluble proteins

His-FGF6, His-FGF8, His-FGF22, His-FGF17, His-FGF16 and His-FGF20 were all expressed, albeit at different levels. After centrifugation, bands corresponding to the molecular sizes of these proteins were detected in the pellet (Fig. 4, compare lanes P and S, green arrow). Although small amounts of protein, such as bands corresponding to His-FGF6, His-FGF16 and His-FGF20, were observed in the supernatants (Fig. 4, lanes S), no protein were detected in the eluate from heparin chromatography, which might suggest these proteins were either small soluble aggregates or not properly folded. It has

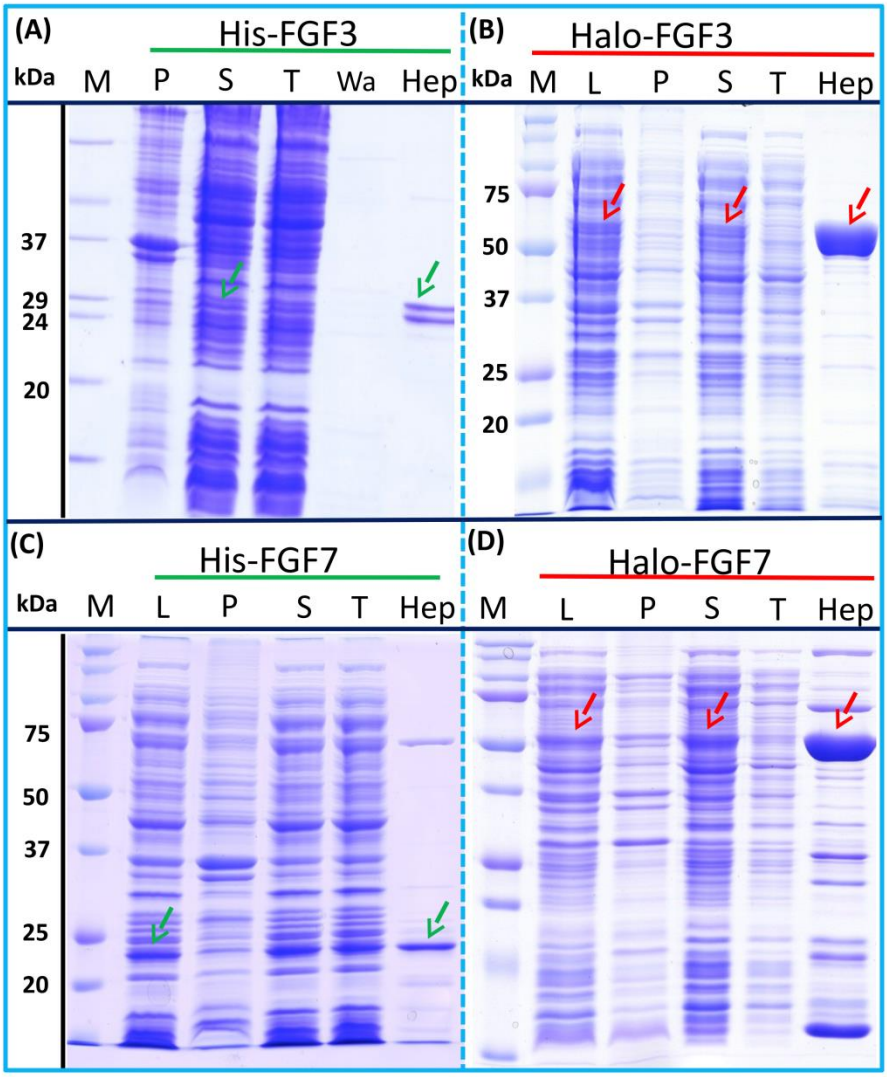


Figure 3 Expression and heparin binding-affinity chromatography of His-FGF3, His-FGF7, Halo-FGF3 and Halo-FGF7. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The supernatant was subjected to heparin-afinity chromatography and samples were then analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant; T, unbound, flow-through fraction from heparin-affinity chromatography; Hep, high [NaCl] eluate of heparin-affinity column (Table 2). Green arrows: His-FGF; red arrows: Halo-FGF.

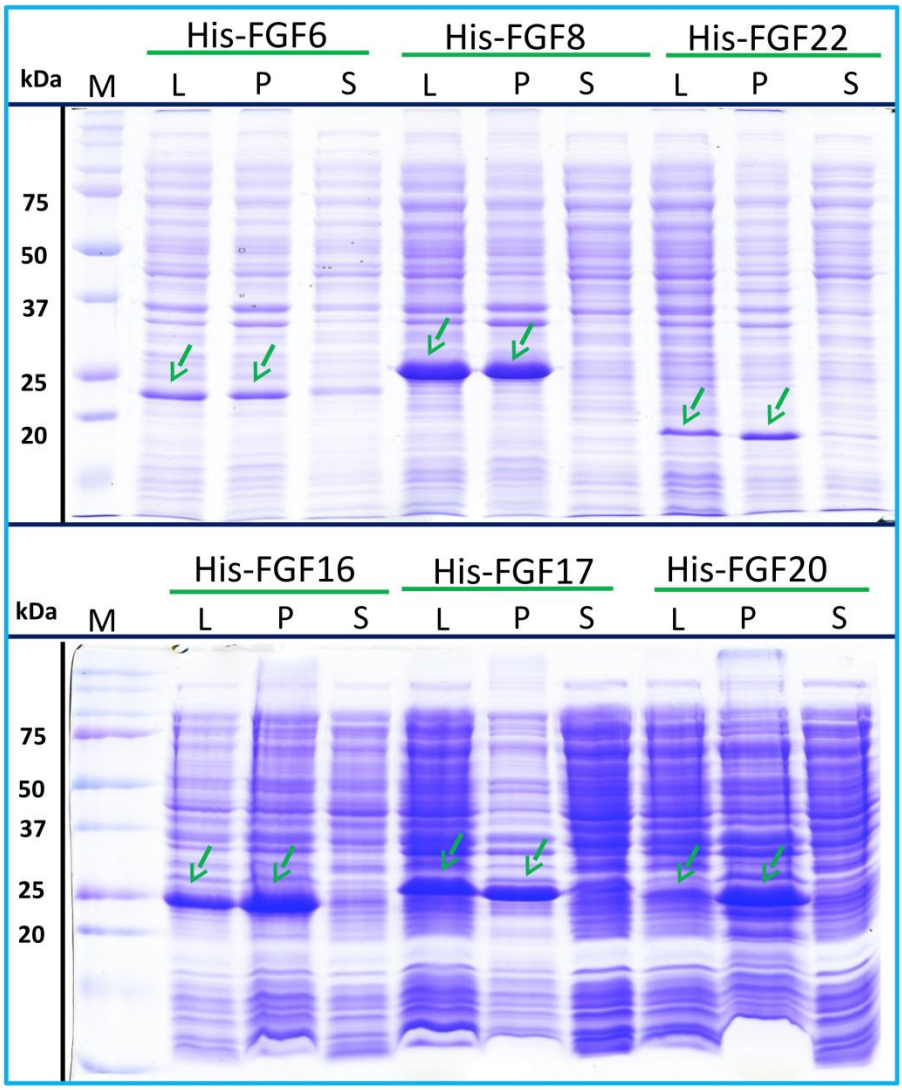


Figure 4 Expression test of His-FGF6, His-FGF8, His-FGF22, His-FGF17, His-FGF16 and His-FGF20. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The whole cell lysate, supernatant and pellet were analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant. Green arrows: His-FGF.

been reported that FGF20 could also be solubilised by high concentrations of arginine (*Maity, Karkaria & Davagnino, 2009*), which suggests that FGF20 in the lysis buffer has a tendency to aggregate. However, arginine would compete for binding of FGFs to heparin, which reduces the utility of this approach to solubilisation.

As illustrated by SDS-PAGE, all of the bands corresponding to the molecular size of Halo-FGF6, Halo-FGF8, Halo-FGF22, Halo-FGF17, Halo-FGF16 and Halo-FGF20 were clearly observed in the whole lysates, which suggested that all six proteins expressed well in *E. coli* (Fig. 5, lanes L, red arrow), particularly Halo-FGF6, Halo-FGF17, Halo-FGF16 and Halo-FGF20. Although some material corresponding to the expected molecular size of these Halo-FGFs was observed in the pellet after centrifugation of the cell lysates (Fig. 5, lanes P), there were strong bands corresponding to Halo-FGF6, Halo-FGF16 and Halo-FGF20 and weak bands corresponding to Halo-FGF8, Halo-FGF17 and Halo-FGF22 present in the soluble fractions (Fig. 5, lanes S). Following application to a heparin affinity column, most of Halo-FGF6 in the supernatant bound to heparin and was eluted by 1 M NaCl in Tris-Cl (Fig. 5A, lanes S, T and Hep). Halo-FGF8 Halo-FGF17 and Halo-FGF22 also bound to the heparin-affinity column reasonably efficiently, whereas a considerable amount of Halo-FGF16 and Halo-FGF20 did not bind (Figs. 5B–5E, lanes S and T). All four proteins could be recovered from heparin chromatography with high concentration NaCl-containing elution buffers (Table 2) (Figs. 5B–5E, lane Hep). When the Halo-FGF20 in the flow-through fraction (Fig. 5F, lane T) was applied to a second heparin-affinity chromatography column, a large amount of Halo-FGF20 was found to bind and could be eluted (Fig. 5F, lane Hep2). A considerable amount of Halo-FGF16 also failed to bind to the heparin affinity column (Fig. 5E, lane T), though the bound protein was eluted with high NaCl (Fig. 5E, lane Hep). This suggests that the capacity of the heparin affinity column for Halo-FGF20 was exceeded. The same explanation may underlie the presence of Halo-FGF16 in the flow-through fraction, though this protein was present at a slightly lower level. However, nothing is known about the preference of either FGF16 or FGF20 for binding structures in the polysaccharide, if so these were relatively rare in heparin, the column capacity might easily be exceeded. Alternatively, the Halo-FGF16 in the flow through fraction may represent protein that is in small aggregates and/or not properly folded.

Given that the amounts of soluble His-tagged FGF6, FGF8, FGF22, FGF17, FGF16 and FGF20 were not readily detectable, it is clear that the N-terminal HaloTag fusion significantly improved the expression of soluble protein. The yield of Halo-FGF6 and Halo-FGF20 was substantial (27 mg/L and 10 mg/L, respectively, Table 3). Although a lower yield of Halo-FGF8, Halo-FGF16, Halo-FGF17 and Halo-FGF22 (1 mg/L to 2 mg/L, Table 3) was obtained, it is nevertheless sufficient for many applications, including microscopy. However, the heparin affinity purification step did not produce entirely pure protein, as judged by coomassie staining (Figs. 2, 3 and 5).

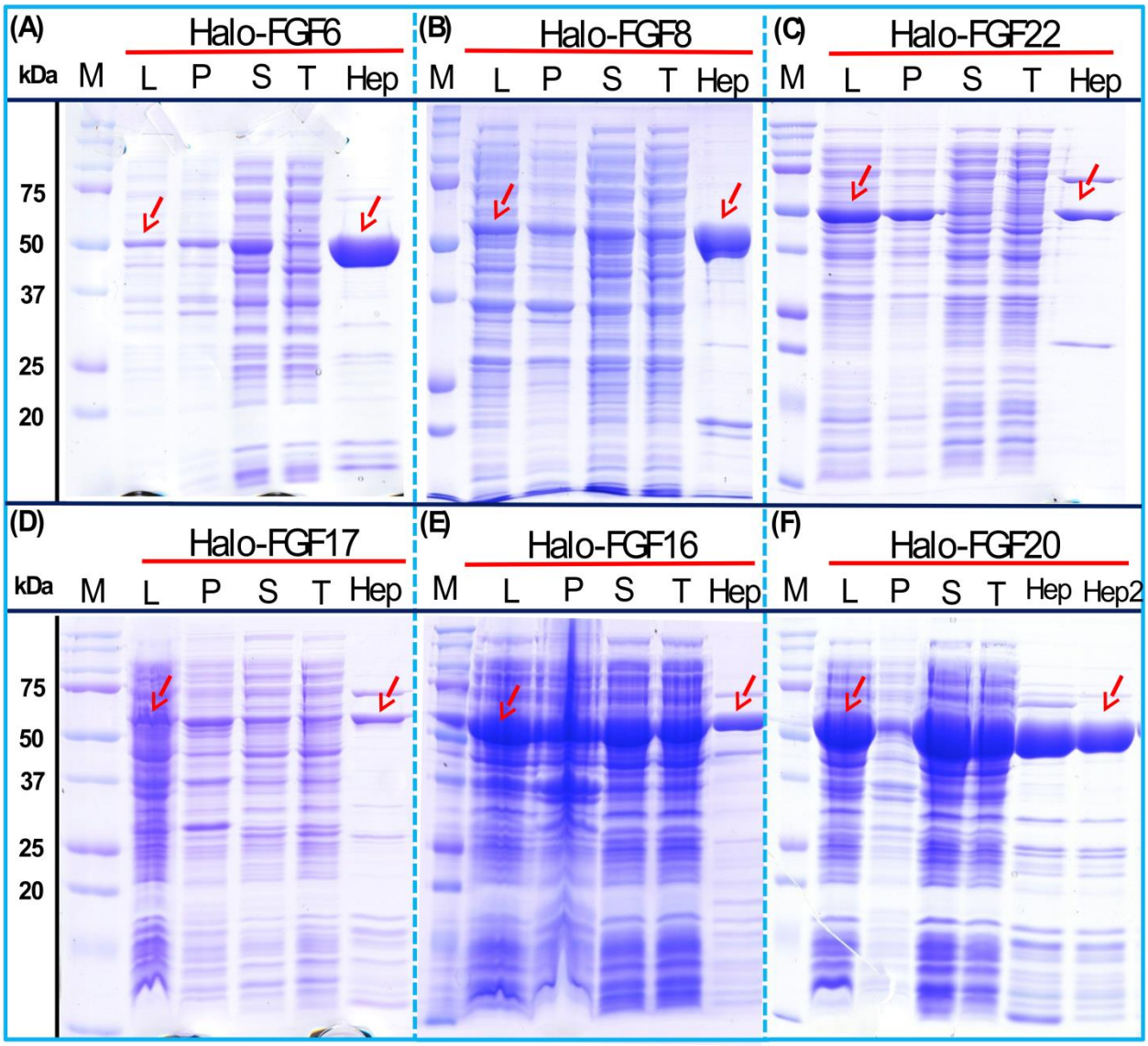


Figure 5 Expression and heparin affinity purification of Halo-FGF6, Halo-FGF8, Halo-FGF22, Halo-FGF17, Halo-FGF16 and Halo-FGF20. Following induction of expression with IPTG, cells were lysed by sonication and the insoluble material collected by centrifugation. The supernatant was subjected to heparin-afinity chromatography and samples were then analysed by SDS-PAGE and coomassie staining. Lane M, markers; L, sonicated whole cell lysate; P, pellet following centrifugation of lysate; S, corresponding supernatant; T, unbound, flow-through fraction from heparin-affinity chromatography; Hep, high NaCl eluate of heparin-affinity column (Table 2). Red arrows: Halo-FGF.

Further purification of some Halo-FGFs

Four Halo-FGFs, Halo-FGF1, Halo-FGF7, Halo-FGF6 and Halo-FGF20 were chosen to determine whether the Halo-FGFs could be easily subjected to further purification, since there was clear evidence for impurities following heparin-affinity chromatography. Halo-FGF1 and Halo-FGF7 were successfully purified by Q anion-exchange chromatography (Figs. 6A and 6B, lane Q), which depends on the acidic isoelectric point of the HaloTag (pI: 4.77). For Halo-FGF6 and Halo-FGF20, advantage was taken of the acidic HaloTag and positive surfaces of FGFs, to enable a two-step ion-exchange purification of the eluate from heparin-affinity chromatography, using both DEAE anion and CM cation ion-exchange chromatography (Figs. 6C and 6D, lane DEAE and CM). The isolated Halo-FGFs are relatively pure, as is shown on the gels (Fig. 6).

Purification of FGFs by removing the HaloTag with TEV protease

The inclusion of a TEV site between the sequence of the HaloTag and FGF proteins provides a means to remove the HaloTag fusion partner in those instances where the HaloTag is not required for analysis (or when it may interfere with such analyses). Halo-FGF2 was first incubated with TEV protease to test whether the fusion protein could be cleaved by TEV. SDS-PAGE of the TEV digestion product of Halo-FGF2 shows that almost all of the protein was cleaved into the 35 kDa HaloTag (Fig. 7A, red arrow) and the 18 kDa FGF2 (Fig. 7A, green arrow). Thus, the cleavage site is fully accessible to TEV protease. Both Halo-FGF17 and Halo-FGF20 were also well digested by TEV protease and subsequently soluble FGF17 (Fig. 7B, green arrow) and FGF20 (Fig. 7C, green arrow) were purified by cation-exchange and heparin chromatography, respectively.

Most of FGF6 (Fig. 7D, lane W_{Dig}, green arrow) and FGF22 (Fig. 7F, lane W_{Dig}, green arrow) and a small proportion of FGF8 were also released from HaloTag (Figs. 7D–7F, lane W_{Dig} and S, red arrow), but these proteins were observed to aggregate upon cleavage. This suggested that these proteins were not very stable, at least in the buffer conditions used here, and required the HaloTag N-terminal fusion to remain soluble. The soluble FGF6 released by cleavage (Fig. 7D, lane S, green arrow) was applied to a heparin affinity column, but was observed to be concentrated at the top of the column where it formed a white aggregate. Very little protein was eluted with 1 M NaCl in PB buffer (Fig. 7D, lane E, green arrow). The disappearance of FGF8 and FGF22 in the soluble fractions after TEV digestion (Figs. 7E and 7F, lane S) showed that these two proteins were also not very soluble in the present buffer conditions without the HaloTag fusion partner.

Biological activities of FGFs and Halo-FGFs on Rama 27 fibroblasts and HaCaT keratinocytes

FGF1, FGF2 and FGF6 have a preference for FGFR1c (*Zhang et al., 2006a*), the predominant receptor expressed by Rama 27 fibroblasts (*Delehedde et al., 2000*; *Zhu et al., 2010*). When Rama 27 cells were stimulated with 25 pM FGF2 for 15 min, strong bands corresponding to dually phosphorylated $p44/42^{MAPK}$ were apparent (Fig. 8A), as observed previously (*Delehedde et al., 2000*; *Zhu et al., 2010*). Halo-FGF2 caused a similar

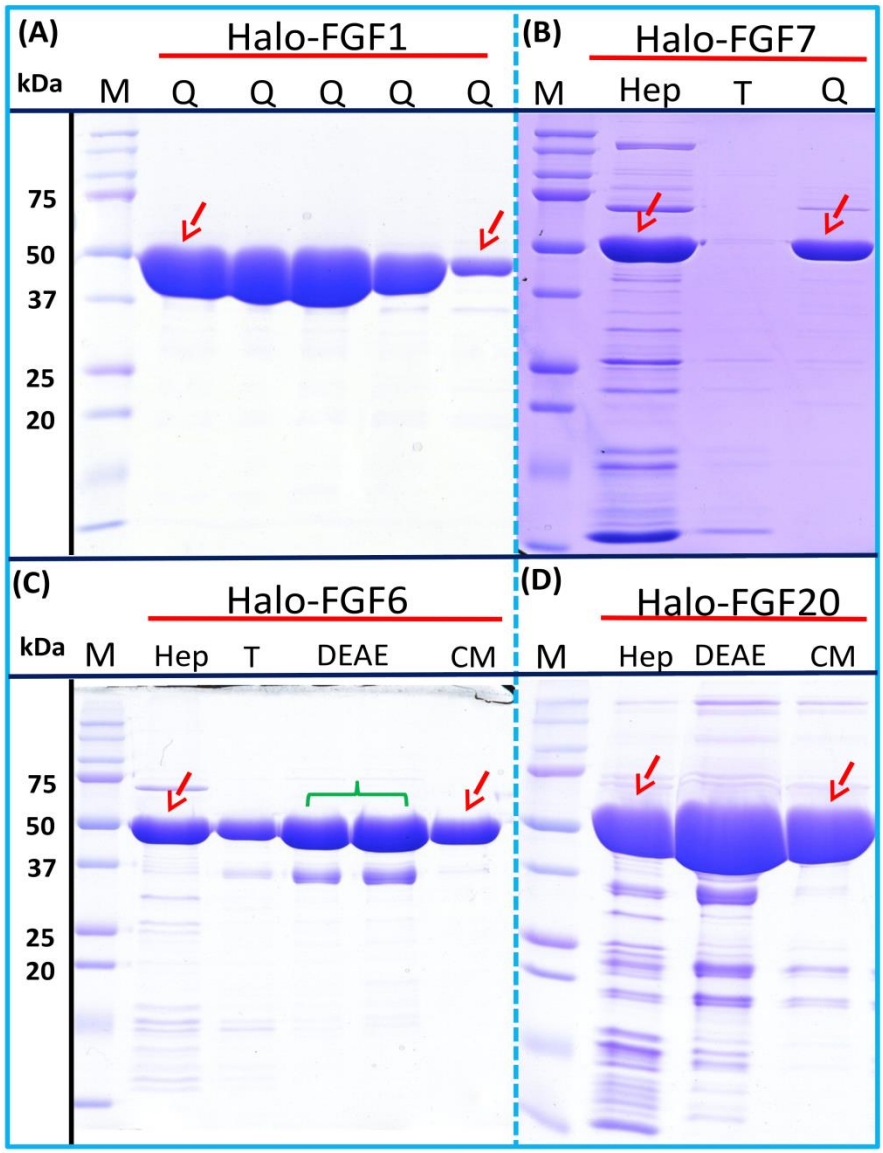


Figure 6 Further purification of the heparin affinity eluate of Halo-FGF1, Halo-FGF6, Halo-FGF7 and Halo-FGF20 by ion-exchange chromatography. The soluble Halo-FGF1 and Halo-FGF7 eluted from heparin chromatography was purified using Q ion-exchange chromatography, while CM and DEAE ion-exchange chromatography were used to purify Halo-FGF6 and Halo-FGF20. Lane M, markers; Hep, eluate from heparin chromatography, Figs. 2A, 3D, 5A and 5F; T, unbound, flow-through fraction from ion-exchange chromatography; Q, peak fractions collected from Q HP chromatography; DEAE eluate from DEAE chromatography, two identical samples; CM, eluate from CM chromatography. Red arrows: Halo-FGF.

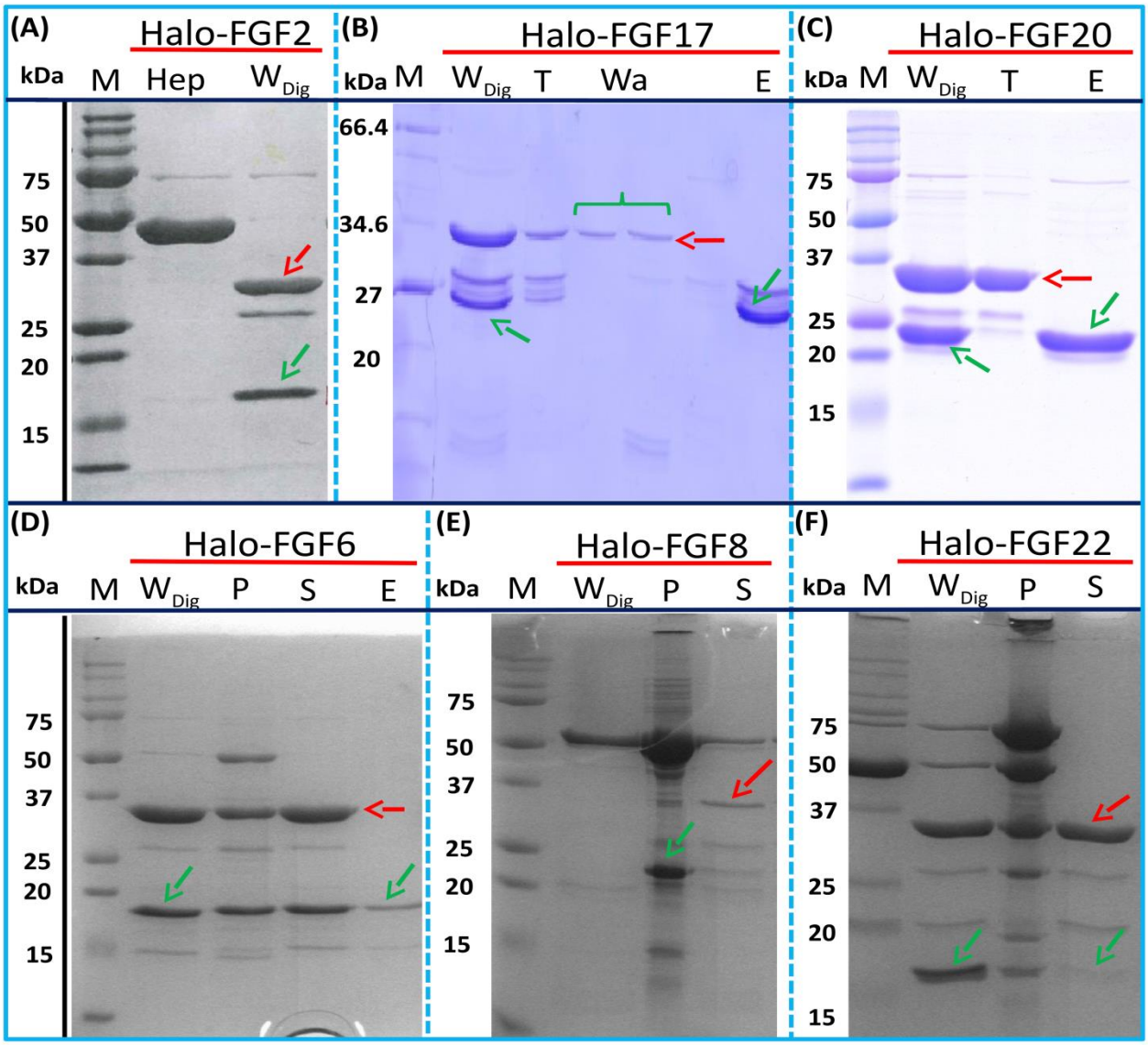


Figure 7 Cleavage of Halo-FGFs by TEV and purification. The eluates of Halo-FGF2, Halo-FGF17, Halo-FGF6, Halo-FGF8 and Halo-FGF22 from heparin-affinity chromatography and the Halo-FGF20 purified by heparin and ion-exchange chromatography were digested by TEV protease to separate the HaloTag and the FGF. Halo-FGF6, Halo-FGF8 and Halo-FGF22 became turbid after digestion and these samples were clarified by centrifugation. Then, the samples containing FGF6 and FGF20 were subjected to heparin chromatography and that of FGF17 to SP HP cation-exchange chromatography. Lanes M, markers; Hep, eluate from heparin chromatography; W_{Dig}, whole digestion product of Halo-FGFs purified by heparin chromatography; T, unbound, flow-through fraction from heparin chromatography; Wa, wash of SP HP cation-exchange chromatography; P, pellet following centrifugation of product of TEV digestion; S, supernatant after the centrifugation; E, high NaCl eluate of heparin or SP cation-exchange chromatography. Green arrows: FGF; red arrows: HaloTag.

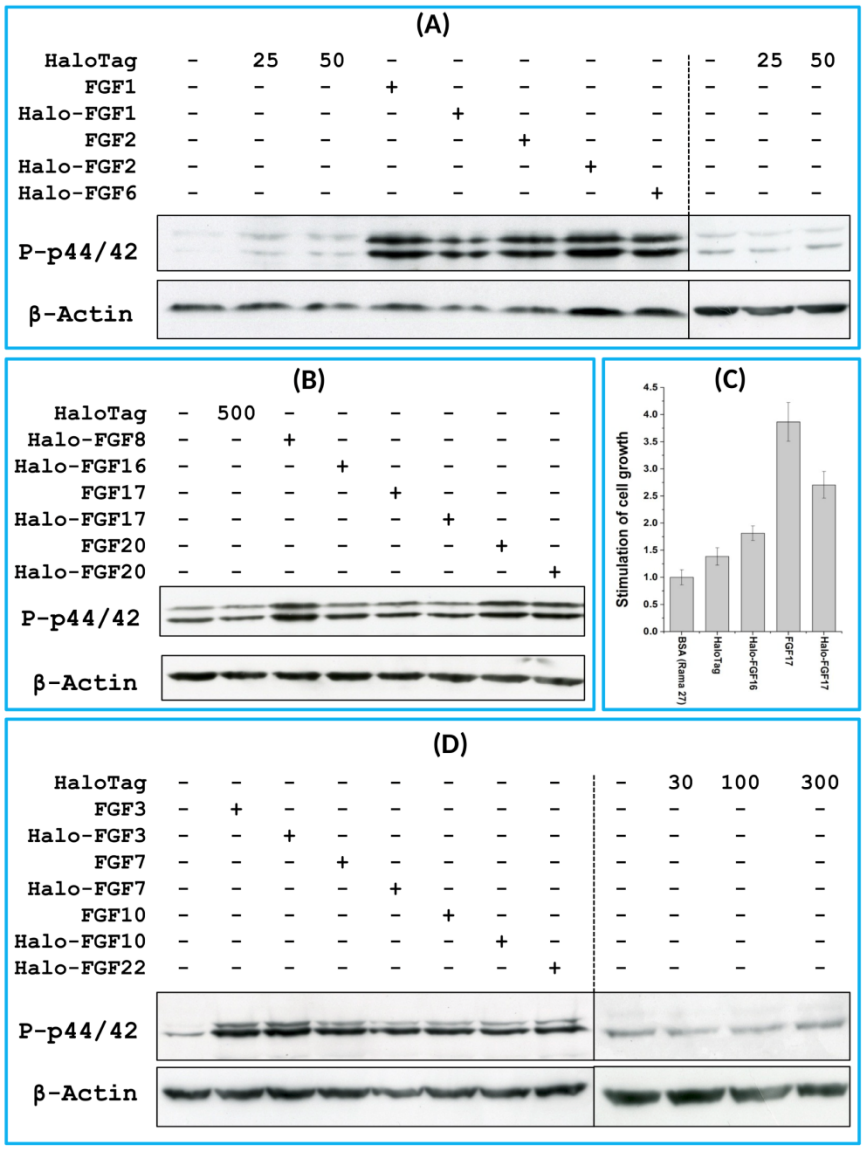


Figure 8 Activities of FGFs on Rama 27 fibroblasts and HaCaT keratinocytes. Cells were grown in 3 cm diameter dishes (Western blots) or 24-well plates growth assay, as described in the Materials and Methods. After incubation in SDM for 24 h (Rama 27) or 48 h (HaCaT), cells were stimulated with the FGF protein for 15 min (Western blot) or 68 h (cell growth assay). (A) Stimulation of p44/42^{MAPK} phosphorylation by 25 pM HaloTag, FGF2, Halo-FGF2 and 50 pM HaloTag, FGF1, Halo-FGF1 and Halo-FGF6 in Rama 27 fibroblasts. (B) Stimulation of p44/42^{MAPK} phosphorylation by 500 pM HaloTag, Halo-FGF8, Halo-FGF16, FGF17, Halo-FGF17, FGF20 and Halo-FGF20 in Rama 27 fibroblasts. (C) Stimulation of cell growth of Rama 27 fibroblasts by 10 nM Halo-FGF16, FGF17 and Halo-FGF17. (D) Stimulation of p44/42^{MAPK} phosphorylation by 300 pM His-FGF3, Halo-FGF3 and Halo-FGF22, 30 pM His-FGF7, Halo-FGF7, His-FGF10 and Halo-FGF10 in HaCaT cells.

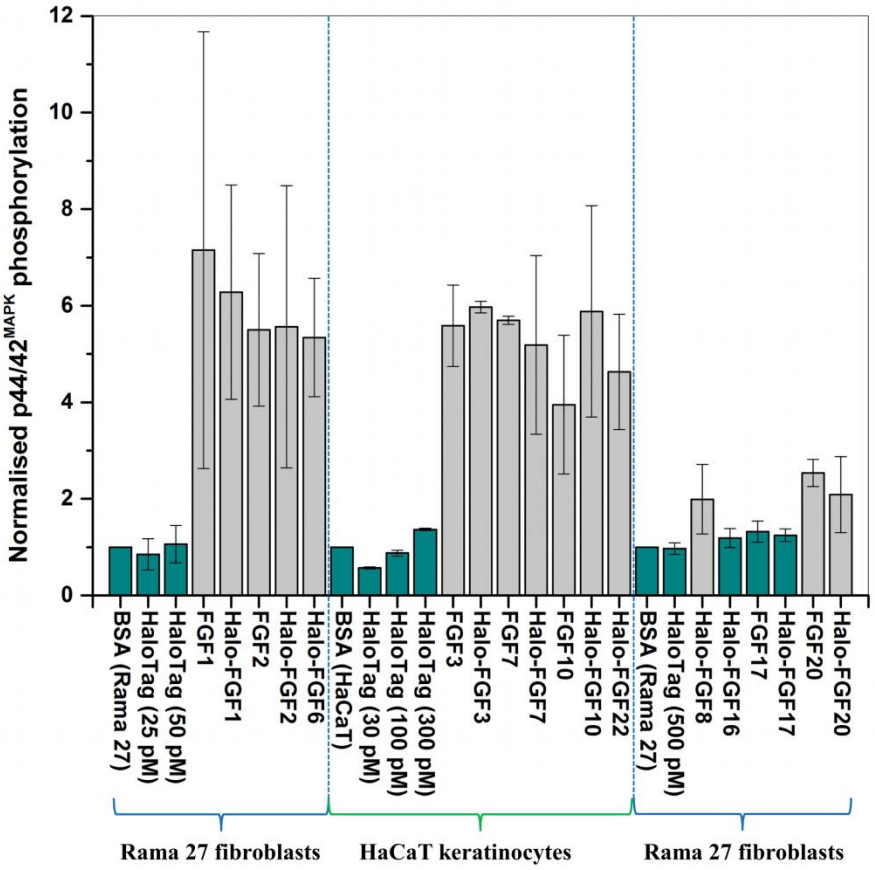


Figure 9 Quantification of p44/42^{MAPK} phosphorylation. The band intensities from two experiments were quantified with imageJ and normalised to the BSA control to compare the similarities and differences of stimulation of phosphorylation p44/42^{MAPK} by different FGFs. Results are the mean with the actual values from two independent experiments.

stimulation of phosphorylation of p44/42^{MAPK} (Fig. 8A). In contrast, the 25 pM or 50 pM HaloTag protein alone did not appreciably stimulate p44/42^{MAPK} phosphorylation. Therefore, the activity of Halo-FGF2 in this assay is equivalent to that of FGF2 (Fig. 9). In the case of FGF1, the N-terminal HaloTag also did not affect the ability of the growth factor to stimulate the phosphorylation of p44/42^{MAPK} (Fig. 8A). FGF6 is not soluble without the HaloTag, so only the activity of the fusion protein could be tested, and it was found to stimulate the phosphorylation of p44/42^{MAPK} to an extent similar to that observed with FGF1 and FGF2 (Fig. 8A). Since FGF6 has the same receptor preference as FGF1 and FGF2 (*Zhang et al., 2006a*), this suggests Halo-FGF6 was fully active.

FGF8, FGF16, FGF17 and FGF20 have a preference for FGFR3c, but they are also able to activate FGFR1c, though higher concentrations of growth factor are required to elicit

activity (*Zhang et al., 2006a*). When 500 pM HaloTag was added to the cells, there was no detectable increase in phosphorylation of p44/42^{MAPK}, whereas Halo-FGF8, Halo-FGF20 and FGF20 at concentrations comparable to those used in previous work (*Zhang et al., 2006a*) were all found to stimulate the phosphorylation of p44/42^{MAPK} (Fig. 8B). In contrast, Halo-FGF16, Halo-FGF17 and FGF17 did not cause a detectable increase in phosphorylation of p44/42^{MAPK} (Fig. 9). These data indicate that Halo-FGF8, FGF20 and Halo-FGF20 have biological activities on Rama 27 fibroblasts. The absence of stimulation of phosphorylation of p44/42^{MAPK} by Halo-FGF16 may reflect the fact that the ability of this FGF to activate FGFR1c is considerably lower than that of FGF8, FGF17 and FGF20 (*Zhang et al., 2006a*). However, the absence of stimulation of phosphorylation of p44/42^{MAPK} by FGF17 and Halo-FGF17 is more puzzling. One explanation may be that FGF16, and perhaps FGF17, do not cause the FGFR to activate strongly early biochemical signals that converge on p44/42^{MAPK}. To test this, the capacity of Halo-FGF16, Halo-FGF17 and FGF17 to stimulate cell growth was measured in Rama 27 fibroblasts. The results show that 10 nM HaloTag only weakly stimulated the growth of Rama27 fibroblasts. Halo-FGF16 caused the number of cells to double compared to the negative control, and this level was significantly ($p = 0.015$, Tukey test, OriginPro 9) above that observed in the presence of HaloTag alone (Fig. 8C). Halo-FGF17 and FGF17 were even more effective, as they caused a 3- to 4-fold increase in the number of cells (Fig. 8C). These results demonstrated that Halo-FGF16, FGF17 and Halo-FGF17 possess biological activities of similar potency as observed by others in growth assays (*Zhang et al., 2006a*).

The activity of members of the FGF7 subfamily were tested on HaCaT keratinocytes, as this cell type expresses the cognate receptor for these FGFs, FGFR2b (*Ron et al., 1993*). HaCaT cells have previously been shown to express more p42^{MAPk} than p44^{MAPk} (*Delehedde et al., 2002*). The data show clearly that HaloTag alone did not stimulate the phosphorylation of p44/42^{MAPK} (Fig. 8D). In contrast, FGF3, FGF7 and FGF10, and the corresponding HaloTag fusion proteins stimulated p44/42^{MAPK} phosphorylation (Fig. 8C). FGF22, which is only soluble as a HaloTag fusion protein, also stimulated p44/42^{MAPK} phosphorylation to an extent similar to that seen with the other members of the subfamily (Fig. 8D). Thus, these Halo-FGFs retain full biological activity in this assay.

CONCLUSION

In this study, we identified four useful properties of N-terminal HaloTag fusions for the production of biologically active FGFs: (i) using the HaloTag can increase the yield of low expression FGFs, (ii) the HaloTag rendered FGF7 non-toxic; (iii) for the insoluble FGFs, the HaloTag enabled *E.coli* to express more soluble protein at low induction temperatures and maintain solubility during isolation and storage; (iv) a consequence of the low isoelectric point of HaloTag was that anion-exchange chromatography could be used as an orthogonal step in the purification of the Halo-FGFs. However, there are clearly limitations; for example, some of the FGFs did not retain solubility following cleavage from the HaloTag. This may reflect the fact that no single solubilisation tag is a universal panacea for resolving the problems of protein expression (*Costa et al., 2014*). Nevertheless, because

the HaloTag can enhance expression of soluble protein and provide a means to label FGF protein with different fluorescent dyes and quantum dots, e.g., *Los et al. (2008)*; *Zhang et al. (2006b)* it is clearly a versatile and useful tool for these two purposes and, therefore, worthwhile exploring as a part of experimental strategy with these aims.

ACKNOWLEDGEMENT

Xianqing Mao would like to thank Monika Dieterle for help with cloning Halo-FGF3.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The Cancer and Polio Research Fund and North West Cancer Research provided financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Cancer and Polio Research Fund and North West Cancer Research.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Changye Sun and Yong Li conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Sarah E. Taylor and Xianqing Mao performed the experiments, contributed reagents/materials/analysis tools, reviewed drafts of the paper.
- Mark C. Wilkinson performed the experiments, reviewed drafts of the paper.
- David G. Fernig conceived and designed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.

REFERENCES

Asada M, Shinomiya M, Suzuki M, Honda E, Sugimoto R, Ikekita M, Imamura T. 2009. Glycosaminoglycan affinity of the complete fibroblast growth factor family. *Biochimica et Biophysica ACTA/General Subjects* **1790**:40–48 DOI 10.1016/j.bbagen.2008.09.001.

Beenken A, Mohammadi M. 2009. The FGF family: biology, pathophysiology and therapy. *Nature Reviews Drug Discovery* **8**:235–253 DOI 10.1038/nrd2792.

Boukamp P, Petrussevska RT, Breitkreutz D, Hornung J, Markham A, Fusenig NE. 1988. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *Journal of Cell Biology* **106**:761–771 DOI 10.1083/jcb.106.3.761.

Costa S, Almeida A, Castro A, Domingues L. 2014. Fusion tags for protein solubility, purification, and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Frontiers in Microbiology* **5**:63 DOI 10.3389/fmicb.2014.00063.

Danilenko DM, Montestruque S, Philo JS, Li TS, Hill D, Speakman J, Bahru M, Zhang MS, Konishi O, Itoh N, Chirica M, Delaney J, Hernday N, Martin F, Hara S, Talvenheimo J, Narhi LO, Arakawa T. 1999. Recombinant rat fibroblast growth factor-16: structure and biological activity. *Archives of Biochemistry and Biophysics* **361**:34–46 DOI 10.1006/abbi.1998.0967.

Delehedde M, Lyon M, Vidyasagar R, McDonnell TJ, Fernig DG. 2002. Hepatocyte growth factor/scatter factor binds to small heparin-derived oligosaccharides and stimulates the proliferation of human HaCaT keratinocytes. *Journal of Biological Chemistry* **277**:12456–12462 DOI 10.1074/jbc.M111345200.

Delehedde M, Seve M, Sergeant N, Wartelle I, Lyon M, Rudland PS, Fernig DG. 2000. Fibroblast growth factor-2 stimulation of p42/44(MAPK) phosphorylation and I kappa B degradation is regulated by heparan sulfate/heparin in rat mammary fibroblasts. *Journal of Biological Chemistry* **275**:33905–33910 DOI 10.1074/jbc.M005949200.

Duchesne L, Gentili D, Comes-Franchini M, Fernig DG. 2008. Robust ligand shells for biological applications of gold nanoparticles. *Langmuir* **24**:13572–13580 DOI 10.1021/la802876u.

Duchesne L, Octeau V, Bearon RN, Beckett A, Prior IA, Lounis B, Fernig DG. 2012. Transport of fibroblast growth factor 2 in the pericellular matrix is controlled by the spatial distribution of its binding sites in heparan sulfate. *PLoS Biology* **10**:e1001976 DOI 10.1371/journal.pbio.1001361.

Ferreira T, Rasband W. 2012. ImageJ user guide—Analyze: Gels. *Available at http://rsbweb.nih.gov/ij/docs/guide/146-30.html#toc-Subsection-30.13* (accessed 09 December 2014).

Huang ZH, Hwang P, Watson DS, Cao LM, Szoka FC. 2009. Tris-nitrilotriacetic acids of subnanomolar affinity toward hexahistidine tagged molecules. *Bioconjugate Chemistry* **20**:1667–1672 DOI 10.1021/bc900309n.

Itoh N. 2007. The FGF families in humans, mice, and zebrafish: their evolutional processes and roles in development, metabolism, and disease. *Biological & Pharmaceutical Bulletin* **30**:1819–1825 DOI 10.1248/bpb.30.1819.

Jeffers M, McDonald WF, Chillakuru RA, Yang MJ, Nakase H, Deegler LL, Sylander ED, Rittman B, Bendele A, Sartor RB, Lichenstein HS. 2002. A novel human fibroblast growth factor treats experimental intestinal inflammation. *Gastroenterology* **123**:1151–1162 DOI 10.1053/gast.2002.36041.

Kalinina J, Byron SA, Makarenkova HP, Olsen SK, Eliseenkova AV, Larochelle WJ, Dhanabal M, Blais S, Ornitz DM, Day LA, Neubert TA, Pollock PM, Mohammadi M. 2009. Homodimerization controls the fibroblast growth factor 9 subfamily's receptor binding and heparan sulfate-dependent diffusion in the extracellular matrix. *Molecular and Cellular Biology* **29**:4663–4678 DOI 10.1128/MCB.01780-08.

Ke YQ, Fernig DG, Smith JA, Wilkinson MC, Anandappa SY, Rudland PS, Barraclough R. 1990. High level production of human acidic fibroblast growth factor in *Escherichia coli* cells inhibition of DNA synthesis in Rat mammary fibroblasts at high concentrations of growth factor. *Biochemical and Biophysical Research Communications* **171**:963–971 DOI 10.1016/0006-291X(90)90778-L.

Knoerzer W, Binder HP, Schneider K, Gruss P, Mccarthy JEG, Risau W. 1989. Expression of synthetic genes encoding bovine and human basic fibroblast growth-factors (bFGFs) in *escherichia-coli*. *Gene* **75**:21–30 DOI 10.1016/0378-1119(89)90379-X.

Lata S, Reichel A, Brock R, Tampe R, Piehler J. 2005. High-affinity adaptors for switchable recognition of histidine-tagged proteins. *Journal of the American Chemical Society* **127**:10205–10215 DOI 10.1021/ja050690c.

Lin XH. 2004. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development* **131**:6009–6021 DOI 10.1242/dev.01522.

Loo BM, Salmivirta M. 2002. Heparin/heparan sulfate domains in binding and signaling of fibroblast growth factor 8b. *Journal of Biological Chemistry* **277**:32616–32623 DOI 10.1074/jbc.M204961200.

Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohane RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J, Darzins A, Klaubert DH, Bulleit RF, Wood KV. 2008. HatoTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology* **3**:373–382 DOI 10.1021/cb800025k.

Macarthur CA, Lawshe A, Xu JS, Santosocampo S, Heikinheimo M, Chellaiah AT, Ornitz DM. 1995. Fgf-8 isoforms activate receptor splice forms that are expressed in mesenchymal regions of mouse development. *Development* **121**:3603–3613.

Maity H, Karkaria C, Davagnino J. 2009. Effects of pH and arginine on the solubility and stability of a therapeutic protein (fibroblast growth factor 20): relationship between solubility and stability. *Current Pharmaceutical Biotechnology* **10**:609–625 DOI 10.2174/138920109789069297.

Ohana RF, Encell LP, Zhao K, Simpson D, Slater MR, Urh M, Wood KV. 2009. HaloTag7: a genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification. *Protein Expression and Purification* **68**:110–120 DOI 10.1016/j.pep.2009.05.010.

Ornitz DM. 2000. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays* **22**:108–112 DOI 10.1002/(SICI)1521-1878(200002)22:2<108::AID-BIES2>3.0.CO;2-M.

Pizette S, Batoz M, Prats H, Birnbaum D, Coulier F. 1991. Production and functional-characterization of human recombinant FGF-6 protein. *Cell Growth and Differentiation* **2**:561–566.

Ron D, Bottaro DP, Finch PW, Morris D, Rubin JS, Aaronson SA. 1993. Expression of biologically-active recombinant keratinocyte growth-factor—structure-function analysis of amino-terminal truncation mutants. *Journal of Biological Chemistry* **268**:2984–2988.

Roullier V, Clarke S, You C, Pinaud F, Gouzer G, Schaible D, Marchi-Artzner V, Piehler J, Dahan M. 2009. High-affinity labeling and tracking of individual histidine-tagged proteins in live cells using Ni^{2+} Tris-nitrilotriacetic acid quantum dot conjugates. *Nano Letters* **9**:1228–1234 DOI 10.1021/nl9001298.

Rudland PS, Twiston Davies AC, Tsao SW. 1984. Rat mammary preadipocytes in culture produce a trophic agent for mammary epithelia-prostaglandin E2. *Journal of Cellular Physiology* **120**:364–376 DOI 10.1002/jcp.1041200315.

Smith JA, Winslow DP, Rudland PS. 1984. Different growth factors stimulate cell division of rat mammary epithelial, myoepithelial and stromal cell lines in culture. *Journal of Cellular Physiology* **119**:120–126 DOI 10.1002/jcp.1041190310.

Susumu K, Medintz IL, Delehanty JB, Boeneman K, Mattoussi H. 2010. Modification of poly(ethylene glycol)-capped quantum dots with nickel nitrilotriacetic acid and self-assembly with histidine-tagged proteins. *Journal of Physical Chemistry C* **114**:13526–13531 DOI 10.1021/jp103872j.

Tinazli A, Tang JL, Valiokas R, Picuric S, Lata S, Piehler J, Liedberg B, Tampe R. 2005. High-affinity chelator thiols for switchable and oriented immobilization of histidine-tagged proteins: a generic platform for protein chip technologies. *Chemistry-a European Journal* **11**:5249–5259 DOI 10.1002/chem.200500154.

Turner N, Grose R. 2010. Fibroblast growth factor signalling: from development to cancer. *Nature Reviews Cancer* **10**:116–129 DOI 10.1038/nrc2780.

Uchinomiya S, Nonaka H, Fujishima S, Tsukiji S, Ojida A, Hamachi I. 2009. Site-specific covalent labeling of His-tag fused proteins with a reactive Ni(II)-NTA probe. *Chemical Communications* **39**:5880–5882 DOI 10.1039/b912025d.

Vogel A, Rodriguez C, IzpisuaBelmonte JC. 1996. Involvement of FGF-8 in initiation, outgrowth and patterning of the vertebrate limb. *Development* **122**:1737–1750.

Xu RY, Ori A, Rudd TR, Uniewicz KA, Ahmed YA, Guimond SE, Skidmore MA, Siligardi G, Yates EA, Fernig DG. 2012. Diversification of the structural determinants of fibroblast growth factor-heparin interactions implications for binding specificity. *Journal of Biological Chemistry* **287**:40061–40073 DOI 10.1074/jbc.M112.398826.

Yu S, Burkhardt M, Nowak M, Ries J, Petrásek Z, Scholpp S, Schwille P, Brand M. 2009. FGF8 morphogen gradient is formed by a source–sink mechanism with freely diffusing molecules. *Nature* **461**:533–536 DOI 10.1038/nature08391.

Zhang XQ, Ibrahimi OA, Olsen SK, Umemori H, Mohammadi M, Ornitz DM. 2006a. Receptor specificity of the fibroblast growth factor family— the complete mammalian FGF family. *Journal of Biological Chemistry* **281**:15694–15700 DOI 10.1074/jbc.M601252200.

Zhang Y, So MK, Loening AM, Yao HQ, Gambhir SS, Rao JH. 2006b. HaloTag protein-mediated site-specific conjugation of bioluminescent proteins to quantum dots. *Angewandte Chemie-International Edition* **45**:4936–4940 DOI 10.1002/anie.200601197.

Zhu HY, Duchesne L, Rudland PS, Fernig DG. 2010. The heparan sulfate co-receptor and the concentration of fibroblast growth factor-2 independently elicit different signalling patterns from the fibroblast growth factor receptor. *Cell Communication and Signaling* **8**:14 DOI 10.1186/1478-811X-8-14.

3.3 Supplementary methods and results

3.3.1 Cloning of His-FGF1, -FGF10, -FGF20, -FGF6, -FGF8 and -FGF22

The cDNAs encoding FGF1, FGF10 and FGF20 were amplified by PCR with the corresponding primers (Table 3.1) from Life Technologies. The vector pET-M11 containing an N-terminal hexahistidine tag was amplified in *E. coli* DH5α and extracted (Section 2.1.3). Then, the three FGF cDNAs and pET-M11 were double-digested (Section 2.1.6) with NcoI (all three FGF cDNAs and pET-M11 vector) and EcoRI (FGF1 and pET-M11) / BamHI (FGF10 and pET-M11) / NotI (FGF20 and pET-M11) (Fig. 3.1). The FGF cDNAs and open pET-M11 vector were extracted from agarose gels (Section 2.1.7). The cDNAs encoding FGF1, FGF10 and FGF20 were inserted into the corresponding pET-M11 vector by T4 ligase and subsequently transformed into Top 10 cells. The bacteria were grown on a karnamycin plate and the colonies were screened by PCR. The plasmids obtained from the colonies corresponding to the correct PCR product were sequenced by GATC Biotech.

Table 3.1 Primers for PCR:

FGF	Primer sequence (5'---3')	Cleavage site
FGF-1	**Forward:** TTAAACCATGGTTAATCTGCCTCCCG	NcoI
	Reverse: TGAATTCTTAATCAGAAGAGACTGGCAGG	EcoRI
FGF-10	**Forward:** TTCAGGGCGCCATGGCGCAAGCCCTTGGTCAAGA	NcoI
	Reverse: GCTGCTCGGATCCCTATGAGTGTACCACCATTGG	BamHI
FGF-20	**Forward:** TTCAGGGCGCCATGGCGATGGCTCCCTTAGCCGAA	NcoI
	Reverse: TGGTGCTCGAGTGCGGCCGCTCAAGTGTACATCAGTAGGT	NotI

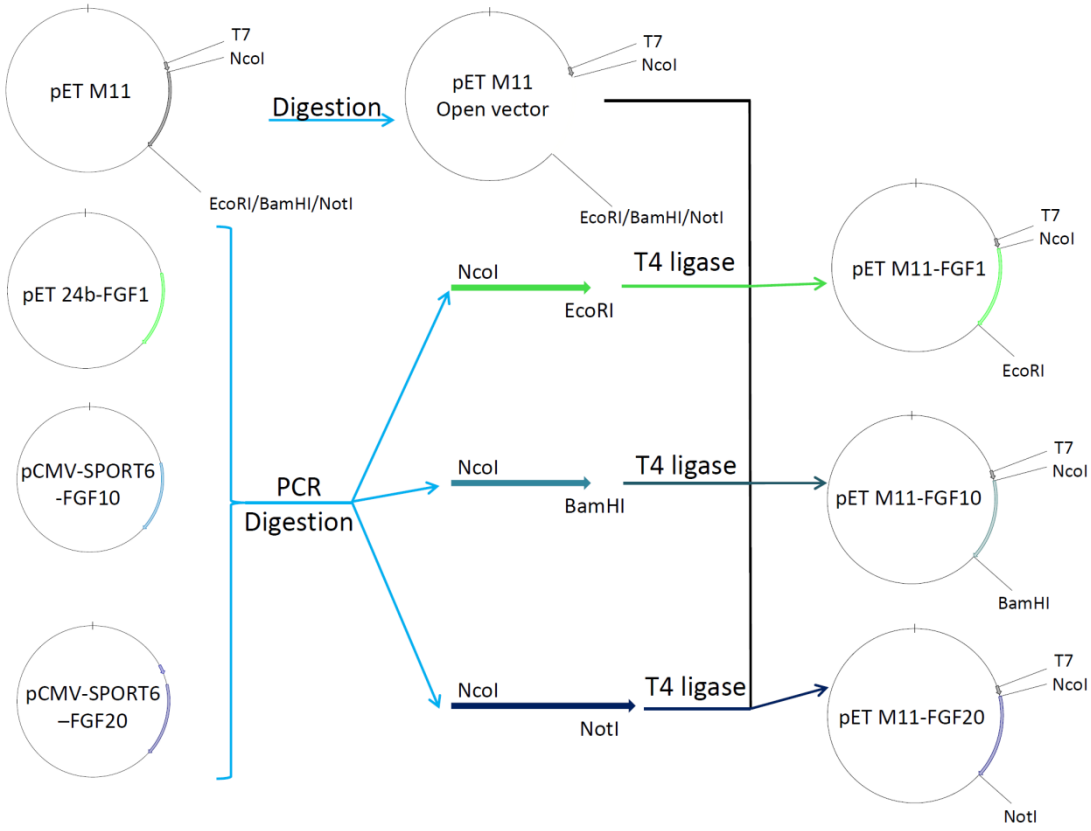


Figure 3.1 Cloning of His-FGF1, 10 and 20. pET M11 vector was prepared by digestion with NcoI and EcoRI/BamHI/NotI restriction enzymes. The same method was used to digest the PCR products of FGF1, 10 and 20. The opened pET M11 vector and DNA fragments encoding FGF1, FGF10 and FGF20 were extracted from agarose gels. FGF cDNAs were inserted into the pET M11 vector using T4 ligase.

The cDNA sequences corresponding to FGF4, 6, 8 and 22 were modified by GeneArt® Gene Synthesis software to enhance expression in *E. coli* and the DNA fragments were synthetized and cloned into the pMA-T vector (Life Technologies). One 5' cleavage site (NcoI) and three 3' cleavage sites (BamHI, NotI and KpnI) were designed into the DNA sequences. The FGF plasmids (pMA-T-FGF6, 8 and 22) and vector (pET-M11) were amplified in *E. coli* DH5α, and were then double digested by NcoI and NotI (Fig. 3.2). The ligation, selection of successfully ligated plasmids and verification by sequencing were accomplished by following the approach used for cloning of His-FGF1, 10 and 20.

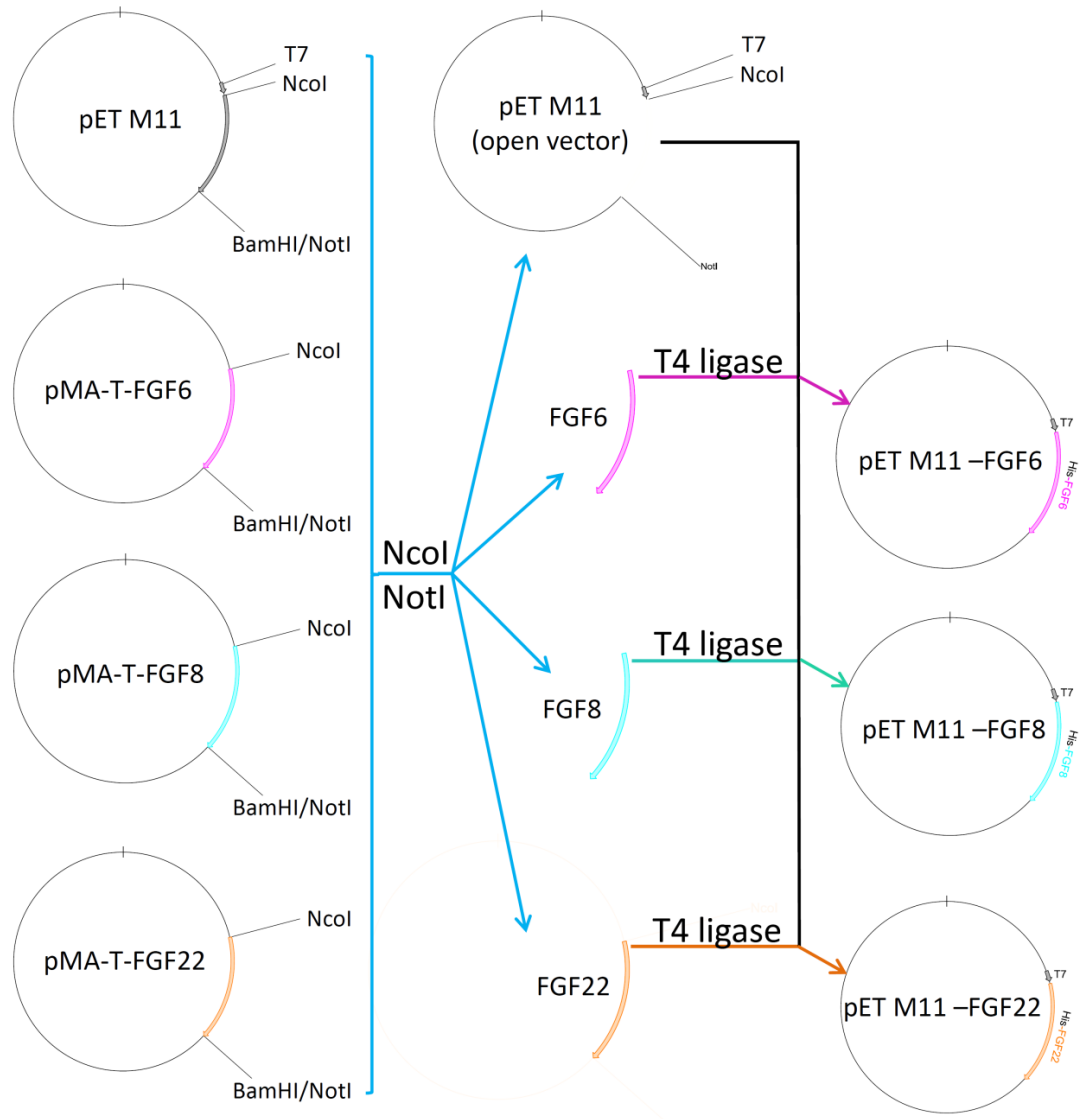


Figure 3.2 Cloning of His-FGF6, 8 and 22. pET M11 vector was prepared by digestion with NcoI and NotI restriction enzymes. The same digestion method was used to extract sequences encoding FGF6, 8 and 22 DNA from the pMA-T vectors. The opened pET M11 vector and DNA fragments encoding FGF6, FGF8 and FGF22 were extracted from DNA agarose gels. DNA encoding FGFs were inserted into pET M11 vector using T4 ligase.

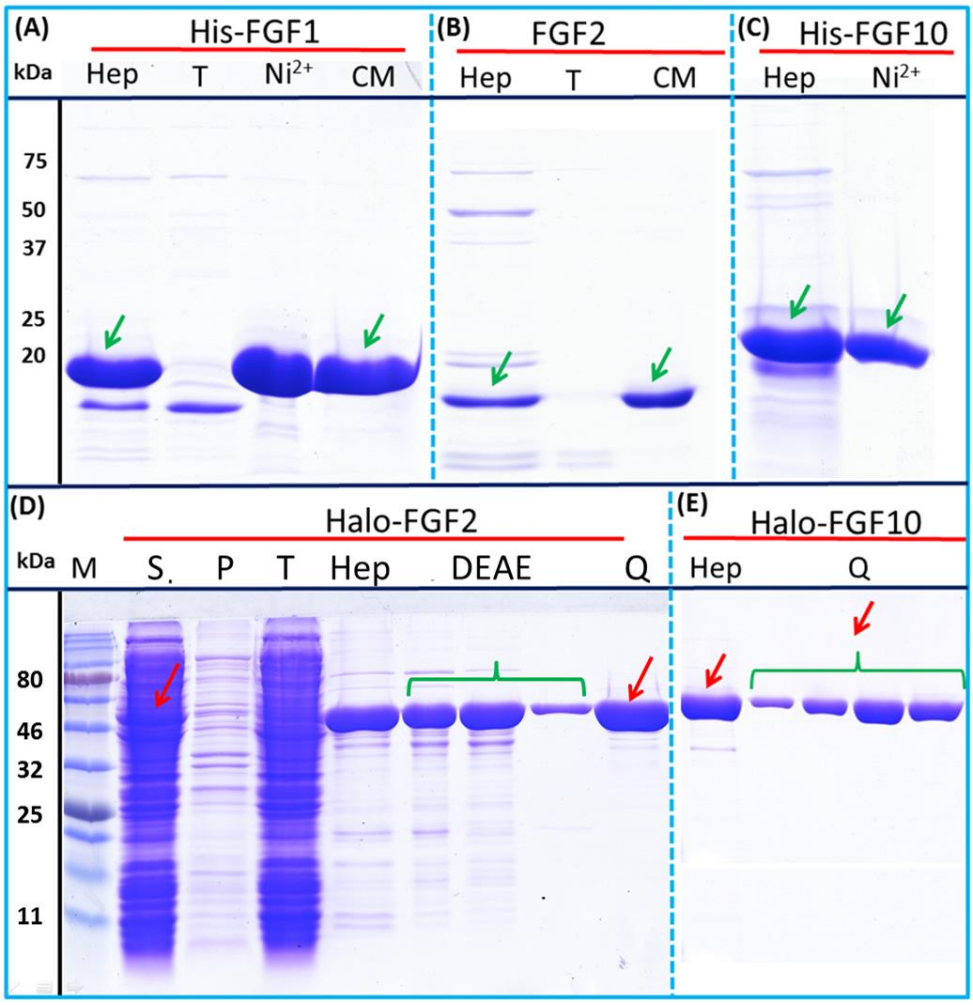


Figure 3.3 Further purification of the heparin-affinity chromatography eluates of Halo-FGF2, Halo-FGF10, His-FGF1, FGF2 and His-FGF10. The soluble Halo-FGF10 eluted from heparin chromatography was purified using Q anion-exchange chromatography, while Halo-FGF2 was purified by DEAE ion-exchange chromatography and Q anion-exchange chromatography. His-FGF1, FGF2 and His-FGF10 eluted from heparin affinity chromatography were further purified by nickel affinity chromatography and/or CM cation-exchange chromatography. Lane M, markers; P, pellet following centrifugation of lysate; S, corresponding supernatant; Hep, eluate from heparin chromatography; T, unbound, flow-through fraction from heparin-affinity chromatography (D) or CM chromatography (B) or nickel chromatography (A); DEAE, eluate from DEAE chromatography, two identical samples; CM, eluate from CM chromatography; Q, peak fractions collected from Q HP chromatography; Red arrows: Halo-FGF, green arrows: His-FGF or FGF2.

3.3.2 Purification of His-FGFs and Halo-FGFs

The His-FGFs and Halo-FGFs were purified by heparin affinity chromatography (Section 3.2). His-FGF1 and His-FGF10 collected from the eluate of the heparin column were further purified by nickel affinity chromatography and washed with 50 mM Tris-Cl (containing 100 mM imidazole and 150 mM NaCl, pH 7.2). The bound His-FGFs were eluted with 50 mM Tris-Cl containing 300 mM imidazole and 150 mM NaCl, pH 7.2 (Figs 3.3 A and C, Ni^{2+} lanes). His-FGF1 from the nickel affinity eluate and FGF2 from the heparin affinity eluate were efficiently purified by CM Sepharose Fast Flow chromatography (Figs 3.3 A and B, CM lanes), as described in Section 3.2. Halo-FGF2 from heparin affinity chromatography was purified by DEAE Sepharose Fast Flow chromatography, but the eluate still contained detectable impurities (Fig. 3.3 D, DEAE lanes). Since the eluate from DEAE chromatography was not sufficiently pure, it was further purified by HiTrap Q HP chromatography, as described in 3.2 (Fig. 3.3 D, Q lane). Halo-FGF10 was also purified by Q HP chromatography and the eluate was relatively pure (Fig. 3.3 E, Q lanes).

3.3.3 Production of SUMO tagged FGF22

Cloning of His-SUMO-FGF22: The cDNA encoding FGF22 was amplified by PCR, and KpnI and HindIII cleavage sites were designed into the PCR product with primers: 5'-CAGGTACCGAGAATCTTTATTTTCAGGGCACCCCGAGCGCAAG TC-3' and 3'-CTACCAAGCTTTTAGCTAACCAGAACC-5'. The POPPINS-3C vector and FGF22 PCR product were double digested by KpnI and HindIII (Figure 3.4). The ligation, selection of successfully ligated plasmids and verification by sequencing of the plasmid encoding His-SUMO-FGF22 were accomplished by following the approach used for cloning of His-FGF1, 10 and 20.

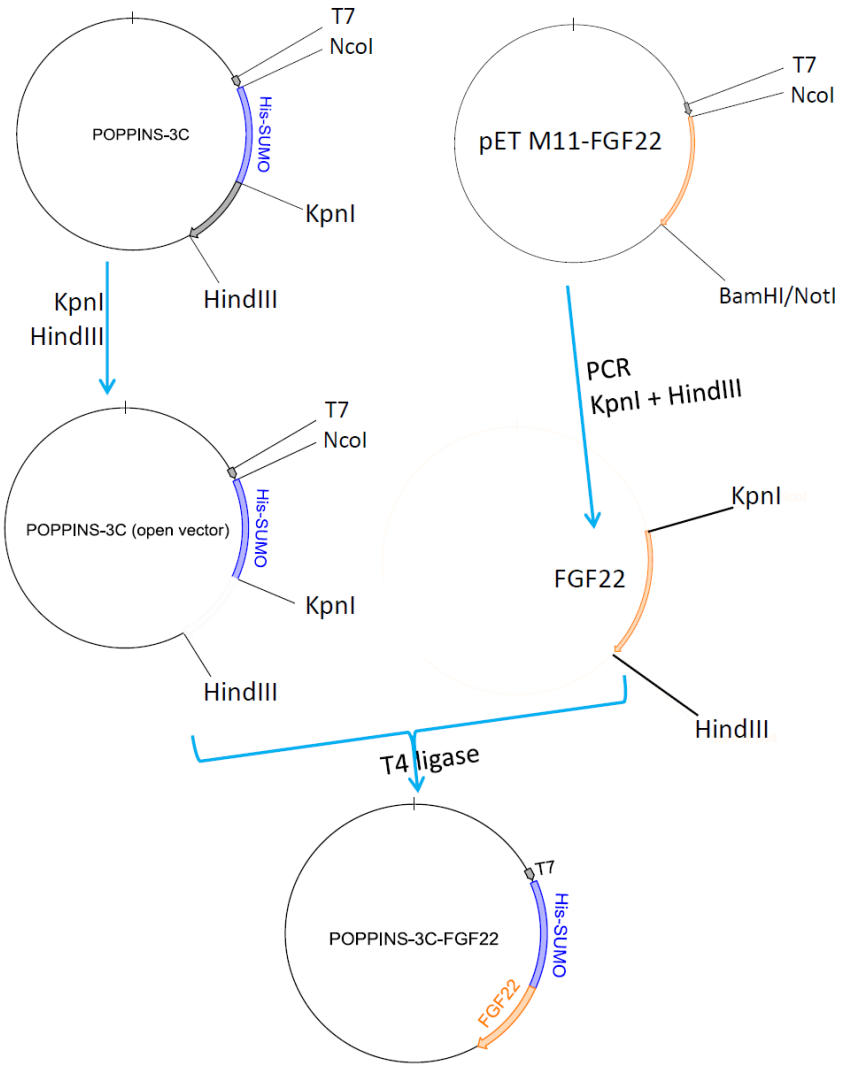


Figure: 3.4 Cloning of His-SUMO-FGF22. POPPIN-3C vector was prepared by digestion with KpnI and HindIII restriction enzymes. FGF22 DNA sequence with 5' KpnI and 3' HindIII cleavage sites was amplified using PCR and was digested with these two enzymes. FGF22 was inserted into POPPIN-3C vector using T4 ligase.

Purification of His-SUMO-FGF22 from soluble fraction: The soluble fraction containing His-SUMO-FGF22 was loaded onto heparin- and nickel-affinity columns, and both the columns were washed with 0.2 M NaCl 50 mM Tris-Cl buffer (pH 7.4). The heparin column was eluted with 2 M NaCl 50 mM Tris-Cl buffer (pH 7.5), and then washed with 8 M urea containing 2 M NaCl. The nickel column was eluted with

300 mM and 600 mM imidazole. The cell lysate was also digested by TEV protease overnight at 4°C to determine the efficacy and effect of cleavage of the SUMO tag. The digestion product was centrifuged at 13000 rpm 4ºC for 10 minutes to separate the soluble and insoluble proteins after digestion. The collected samples were analysed by SDS-PAGE.

Purification result: The molecular weight of His-SUMO-FGF22 is 31.2 kDa. A protein corresponding to the expected size of His-SUMO-FGF22 on SDS-PAGE was produced in the transformed bacteria induced overnight at 16°C by IPTG. This protein was found in the soluble fraction, indicating that the SUMO tag had successful increased the solubility of the FGF22 (Fig. 3.5 A, S lane). However, upon heparin affinity chromatography, most of His-SUMO-FGF22 was found in the flow-through fraction (Fig. 3.5 A, T lane). Similarly, most of His-SUMO-FGF22 failed to bind to the nickel affinity column (Fig. 3.5 A, T' lane). Moreover, the small amount of protein that did bind to the affinity columns could not be eluted by normal elution buffers (Fig. 3.5 A, E-1 and Ni1 lanes), but required 8 M urea with 2 M NaCl (heparin column) or 600 mM imidazole (nickel column). The poor binding of the SUMO-FGF22 to the heparin-affinity column may indicate that the protein might not be correctly folded, since the canonical HBS is composed of residues that are spatially adjacent, but distant in sequence. An alternative explanation is that the protein is correctly folded, which would be consistent with its appearance in the soluble fraction, but was oligomerised such that the canonical HBS and the His tag were not exposed to solvent. That fraction that did bind to the affinity columns is likely to be oligomerised, since it did not all elute from heparin with 2 M NaCl and some required 600 mM imidazole for elution from the nickel affinity column. The digestion product of the fusion protein with TEV protease was not complete and both

the digested (FGF22) and non-digested (His-SUMO-FGF22) fragments were observed in the pellet following centrifugation (Fig. 3.5 B, D-p lane). These data suggest that FGF22 aggregated without the SUMO tag and the incomplete digestion with TEV is consistent with the presence of oligomers with buried N-terminal His tags.

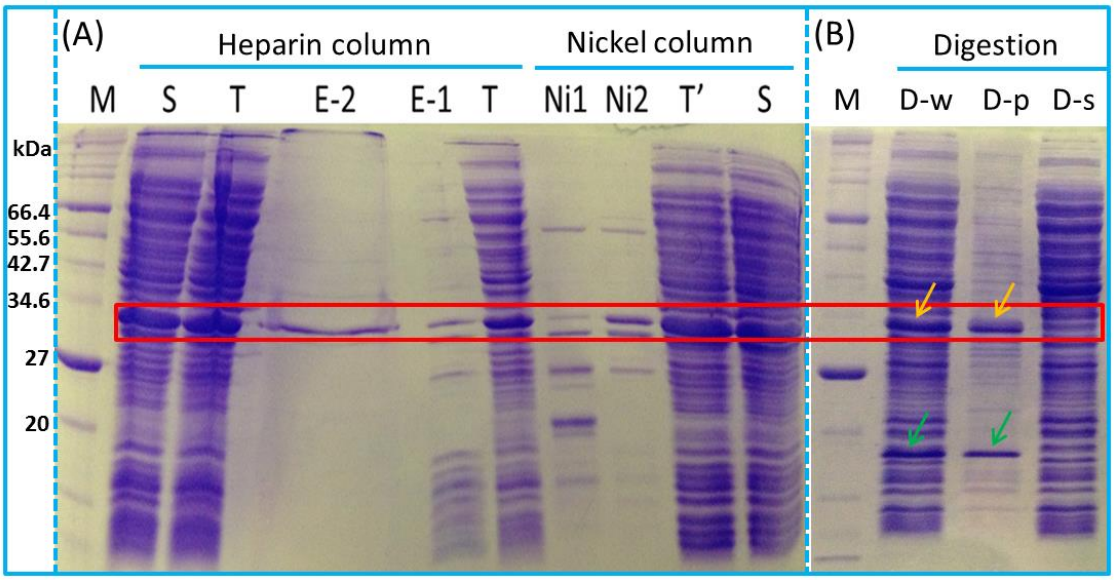


Figure 3.5 Expression and purification of His-SUMO-FGF22. (A): (M), marker (2-212 kDa protein marker, NEB, UK); (S), soluble fraction; (T), flow-through from heparin column; (E-1), eluate from heparin column with 2 M NaCl 50 mM Tris-Cl; (E-2), eluate from heparin column with 2 M NaCl and 8 M urea; (T'), flow-through; (Ni1), eluate with 300 mM imidazole; (Ni2), eluate from with 600 mM imidazole from nickel affinity chromatography; **(B):** (D-w), whole digestion product; (D-p), pellet of digestion product; (D-s), supernatant of digestion product. **Red square:** His-SUMO-FGF22; **Orange arrow:** non-digested His-SUMO-FGF22; **Green arrow:** FGF22.

3.4 Conclusion

Many different FGFs were produced by expression in *E. coli* and then successfully purified. FGF1 and FGF2, which had previously been successfully expressed (Ke et

al., 1990, Ke et al., 1992, Duchesne et al., 2008), were relatively simple to produce as the native or the tagged protein. FGF10, a well-studied member of the FGF7 subfamily (Igarashi et al., 1998), was also straightforward to express and purify. However, the other FGFs were more challenging. Indeed many of these have proved difficult to produce by others (summarised in Section 3.2) and in our laboratory (Xu, 2012).

Table 3.2 Recommended constructs for producing FGFs. The yields of FGFs and Halo-FGFs, their receptor binding specificities (Zhang et al., 2006b) and biological activities are included in this table.

Subfamily	Name	Yield (mg/L)	Receptor specificity	Biological activity
FGF1	FGF1	14	All FGFRs	MAPK
	FGF2	2.5	FGFR 1c, 3c >2c, 1b, 4Δ	MAPK
FGF4	Halo-FGF6	27	FGFR 1c, 2c >3c, 4Δ	MAPK
FGF7	Halo-FGF3	11		MAPK
	Halo-FGF7	5.6	FGFR 2b >1b	MAPK
	FGF10	9.3		MAPK
	Halo-FGF22	2.0		MAPK
FGF8	Halo-FGF8	1.7	FGFR 3c > 4Δ >2c > 1c >> 3b	MAPK
	Halo-FGF17	1.5		Cell growth
FGF9	Halo-FGF16	1.0	FGFR 3c > 2c > 1c, 3b >> 4Δ	Cell growth
	Halo-FGF20	10		MAPK

FGF22 was not soluble (Section 3.2) and initially it was expressed with a SUMO tag to overcome this problem. However, the SUMO tagged FGF22 in the soluble fraction did not bind to heparin, which suggested the SUMO tag might not be able to fully solubilise FGF22. The HaloTag was initially used to provide a means of placing

a fluorescent probe in a structurally defined position on FGF2 (*c.f.* chemical modification of amino acid side chains), and was fortuitously observed to enhance protein expression (Section 3.2). HaloTag was, therefore, explored as a solubilisation tag. It was found to be suitable for the expression of soluble FGF3, FGF6, FGF7, FGF8, FGF17, FGF20 and FGF22, all of which could then be purified (Section 3.2). The recommended constructs for making each FGF are described in Table 3.2. All these proteins possess biological activity based on cell signalling and cell growth assays on cells possessing the appropriate receptors (HaCaT, FGFR2b, activated by FGF7 family members; Rama 27 FGFR1c, activated to different extents by all other FGFs) (Ornitz et al., 1996, Zhang et al., 2006b). In any event, since they bind heparin, they will be properly folded, because the canonical heparin-binding site has contributions from amino acids that are not adjacent in the sequence (Sections 1.5.3 and 1.5.4). Why FGF16 and FGF17 did not elicit a stimulation of phosphorylation of $p44/42^{MAPK}$, but nevertheless stimulated cell growth, is not known. However, the intracellular signalling pathways stimulated by these two FGFs have not been studied, so the present data suggest that their stimulation of cell growth may not involve to a great extent the $p44/42^{MAPK}$ pathway. The yields of some of these Halo-FGFs (Halo-FGF8, 16, 17 and 22) are still low, which suggests HaloTag itself has not solved all the problems for production of soluble FGFs. However, the success of making many FGFs (FGF3, FGF6, FGF7, FGF17 and FGF20) provided enough tools for the analysis of the interaction of these FGFs with their polysaccharide co-receptor, aspects of which are explored in the following chapters.

Chapter 4 Binding specificity of FGFs to heparin/HS

4.1 Introduction

The paracrine FGFs are morphogens involved in development, where they are thought to form different gradients by binding to HS chains in the ECM; subsequently in organism growth and homeostasis, such gradients are also likely to be important (Sections 1.2 and 1.8). Since the structure of HS chains shows some degree of tissue specificity (Section 1.3), then it follows that the ability of different tissues to bind and restrict the movement of different FGFs may also differ. The evolutionary development and natural selection that drove the expansion of these FGFs have been proposed to cause a divergence in their selectivity for binding structures in HS (Xu et al., 2012). Heparin, which can be viewed as a highly sulfated HS, is widely used to purify the paracrine FGFs, which demonstrates that all of the paracrine FGFs can strongly bind to heparin (Asada et al., 2009).

To probe the binding selectivity of FGFs for the glycosaminoglycans, differential scanning fluorimetry (DSF) and a library of chemically modified heparins and model GAGs were used, as described (Uniewicz et al., 2010, Xu et al., 2012). DSF is a thermal shift assay, which follows protein denaturation due to increasing temperature. The melting temperature (Tm) of a protein is often very sensitive to the binding of ligands and buffer conditions. In the case of FGFs, the binding to polysaccharide increases Tm and this change is used as a proxy for affinity. DSF produces a large amount of data that requires analysis. Consequently, user-friendly

software has been developed to enable the efficient analysing and viewing DSF data. In this chapter the experimental procedure, data acquisition and analysis are introduced and the binding selectivity of two FGFs, FGF10 and FGF20 to the library of polysaccharides is measured by DSF and discussed. In previous work, the selectivity of the interactions with glycosaminoglycans of two members of the FGF1 subfamily, along with one member of the FGF7 (FGF7), FGF8 (FGF18) and FGF9 (FGF9) subfamilies was measured (Uniewicz et al., 2010, Xu et al., 2012). A hypothesis generated from these data was that the specificity of the FGF-glycosaminoglycan interactions had been subjected to the same natural selection pressures that led to the diversification of the FGF family (Xu et al., 2012, Xu et al., 2013). An aim of this chapter was to acquire data on FGFs from subfamilies that hitherto only had one member characterised, as this would enable the hypothesis to be tested. To this end, the interactions of FGF10, in the FGF7 subfamily, and FGF20, in the FGF9 subfamily (Section 1.5.1) with glycosaminoglycans were characterised.

4.2 DSF experimental procedure, data acquisition and analysis (method paper)

SimpleDSFviewer: a tool to analyse and view differential scanning fluorimetry data for characterising protein thermal stability and interactions

Changye Sun, Yong Li, Edwin A. Yates, David G. Fernig

Department of biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK

PeerJ PrePrints 3:e1937 (DOI: 10.7287/peerj.preprints.1555v1*)*

Contributions of the authors:

Changye Sun: Produced FGF10; wrote the program code; tested the program with example data; prepared the figures; Conceived study and wrote the paper. **With Yong Li:** Produced FGF7; tested the program with some experimental results; edited the paper. **Ed A. Yates:** Conceived study and edited the paper. **David G. Fernig:** Conceived study and edited the paper.

Key words: differential scanning fluorimetry (DSF), protein melting curve, protein melting temperature, data analysis

Abstract: Differential scanning fluorimetry (DSF) is used widely as a thermal shift assay to study protein stability and protein-ligand interactions. The benefit of DSF is that it is simple, cheap and can generate melting curves in 96 well plates providing good throughput. However, data analysis remains a challenge. In this article, the program, SimpleDSFviewer, is introduced as a user-friendly interface to help view and analyse DSF data.

1 Introduction

Differential scanning fluorimetry (DSF) is a thermal shift assay method used to measure the denaturation of proteins due to increasing temperature. Increased thermal energy will break the non-covalent bonds that underlie protein folding (Niesen et al., 2007, Semisotnov et al., 1991, Pantoliano et al., 2001). A high temperature is required to denature a stable protein, while an unstable protein will be denatured at lower temperature. The fluorescent dye, which gives high fluorescence in a non-polar environment, is used in the DSF assay to probe the hydrophobic sites on the unfolded proteins. So, as the temperature is increased, the protein unfolds to expose more and more hydrophobic residues until the protein is fully unfolded. This in turn produces more fluorescence by the interaction of fluorescent dye and the exposed hydrophobic residues. Since the DSF method has been adopted by an ever wider community, an increasing number of applications have been developed to test protein stability and protein-ligand interactions (Vivoli et al., 2014). The equipment requirement is also so modest - a common RT-PCR instrument and 96 multi-well plates that the experiment can be done easily and a large volume of data can be acquired (Niesen et al., 2007, Vivoli et al., 2014). However, analysis of the data is more challenging. Consequently, a program for DSF data analysis has been written

to help users of DSF view and analyse efficiently their data. The program has a user-friendly interface and a number of additional functions have been included, such as normalisation, smoothing and melting temperature extraction. Worked examples with two proteins, FGF7 and FGF10 are shown. FGF7 serves as a reference, since its interactions with heparin have been previously characterised by DSF (Xu et al., 2012). In contrast, FGF10, which is in the same FGF subfamily, has not had its interactions with heparin characterised previously.

2 Data acquisition

Proteins and ligands – In the given example, the reagents His-FGF10 (hexahistidine tagged Fibroblast Growth Factor 10) and His-FGF7 were produced, as described (Xu et al., 2013, Sun et al., 2015). Heparin sodium powder (from porcine intestinal mucosa, Sigma-Aldrich, Dorset, UK) was dissolved in phosphate-buffered saline (PBS buffer: 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 and 137 mM NaCl, pH 7.4) to prepare a series of concentrations of ligands (supplementary: Example data) for FGF binding.

Differential Scanning Fluorimetry (DSF) setup – The purified FGFs, heparin ligand and SYPRO Orange dye (Life technologies, Paisley, UK) were added into the corresponding buffer and mixed gently, as listed in Table 1 (Uniewicz et al., 2010, Xu et al., 2013). Fast optical 96-well plates (Life technologies) were used to carry samples, which were read by a 7500 Fast Real-time PCR machine (Life technologies), and the prepared samples were split into three wells, 10 µL in each. The plate was then covered with Optical Adhesive Film to prevent evaporation. The running method was designed to raise the temperature from 32°C to 81°C in 0.5°C steps, and to measure the fluorescence every 0.5°C for 30 seconds. The protein concentration used for data acquisition was 5 µM. The raw data were exported from the running software (7500 Fast Real-time PCR system) to an Excel file, after the fluorescence reading. The details of the program for data analysis are described in the **supplementary**.

Table 1 DSF reaction components for three wells:

Components	Protein	Ligand (10X)	SYPRO Orange (100X)	Buffer	Total volume

Volume / µL	a[1]	3.5	3.5	(28-a)	35

a[1]: volume of the protein for specified concentration.

3 Data analysis and display with SimpleDSFviewer using Matlab

3.1 Basic functions for data analysis and view – The SimpleDSFviewer automatically processes the raw data or the re-organised data (Fig. S2), as shown in the flow chart (Fig. 1). The melting curve data are normalised and smoothed with the given range and the first derivative of the melting curve is calculated as described in 3.2. The melting temperature of each melting curve is extracted (3.3). The analysed data may be saved in an Excel file format for further analysis.

The melting curve patterns of the tested proteins are also automatically screened and presented on an image, as is shown in 3.4 and 3.5, which gives a view of the proteins' thermal stabilities across the 96 wells. The protein melting curves and the first derivative curves of selected wells can be plotted in the preview box or in a separate figure, which can be automatically or manually saved in different formats. The melting temperature point for each protein is also marked on the melting curve or first derivative curve.

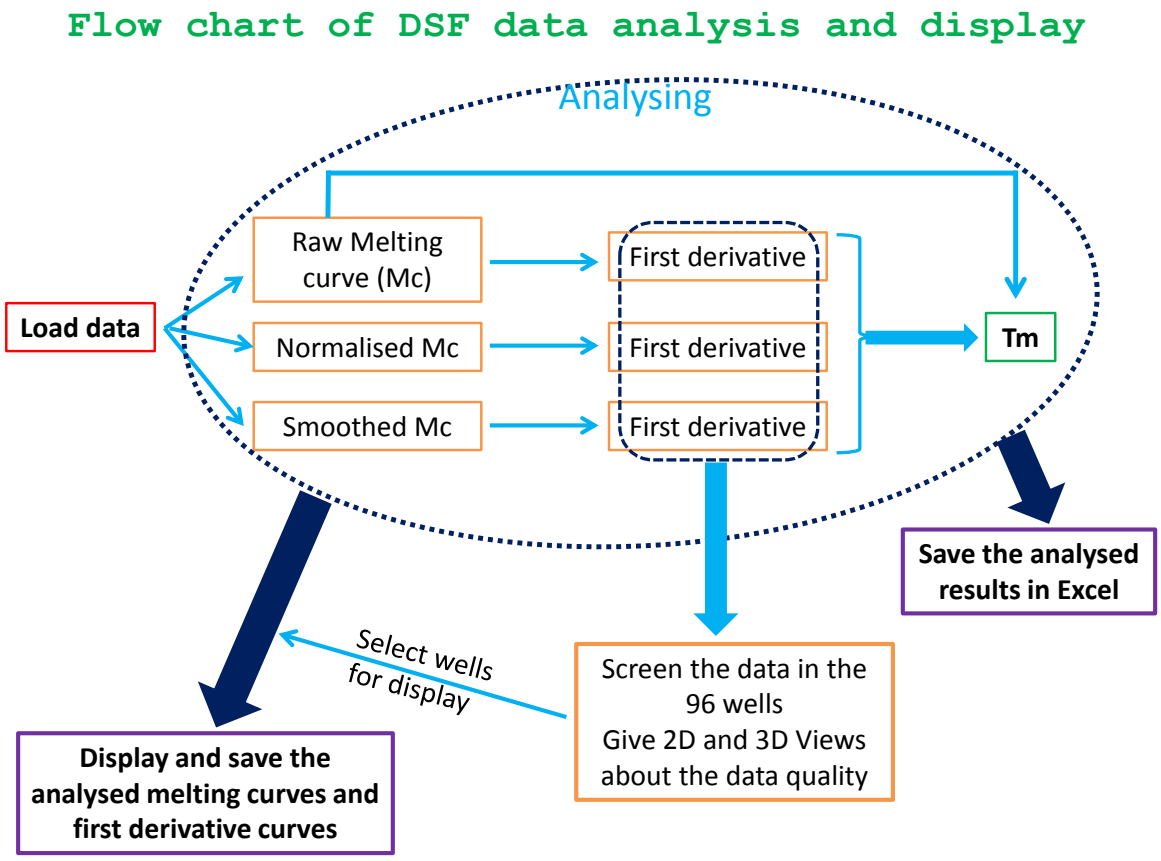


Figure 1 Flow chart of DSF data analysis and display. The processing steps for the input data are described in this diagram. The red box and green boxes are the input (experimental data) and one output (melting temperature). The yellow boxes are the automatically analysed outputs. The purple boxes are to display data and to save data. The 'Screen Data' function directs the user to select the desired data for viewing.

3.2 Normalise, smooth melting curve and first derivative curve – The background is removed by subtracting the minimum value of each curve from all points, and then the melting curve is normalised by dividing by the maximum value (Equation 1). The normalisation of melting curves enables these to be compared to each other (Figs 2 A-B). The first derivative is calculated using Equation 2 and is normalised in the same way as the melting curves. The melting curves may be slightly noisy, *e.g.*, due to low protein concentration (Figs 2 C-D), so each normalised curve is smoothed in three parts, the beginning to the minimum value, the minimum value to the maximum value and the maximum to the end, to avoid smoothing the part containing inflection point and ensure the original shape of the curve is maintained. Due to the

high quality smooth, the raw melting curve is strongly covered by the smooth curve. The smoothing method is based on Matlab - smooth function (MathWorks, 2015) and the smooth range can be defined by the user. Note that the 'normalise' and 'smooth' functions are optional.

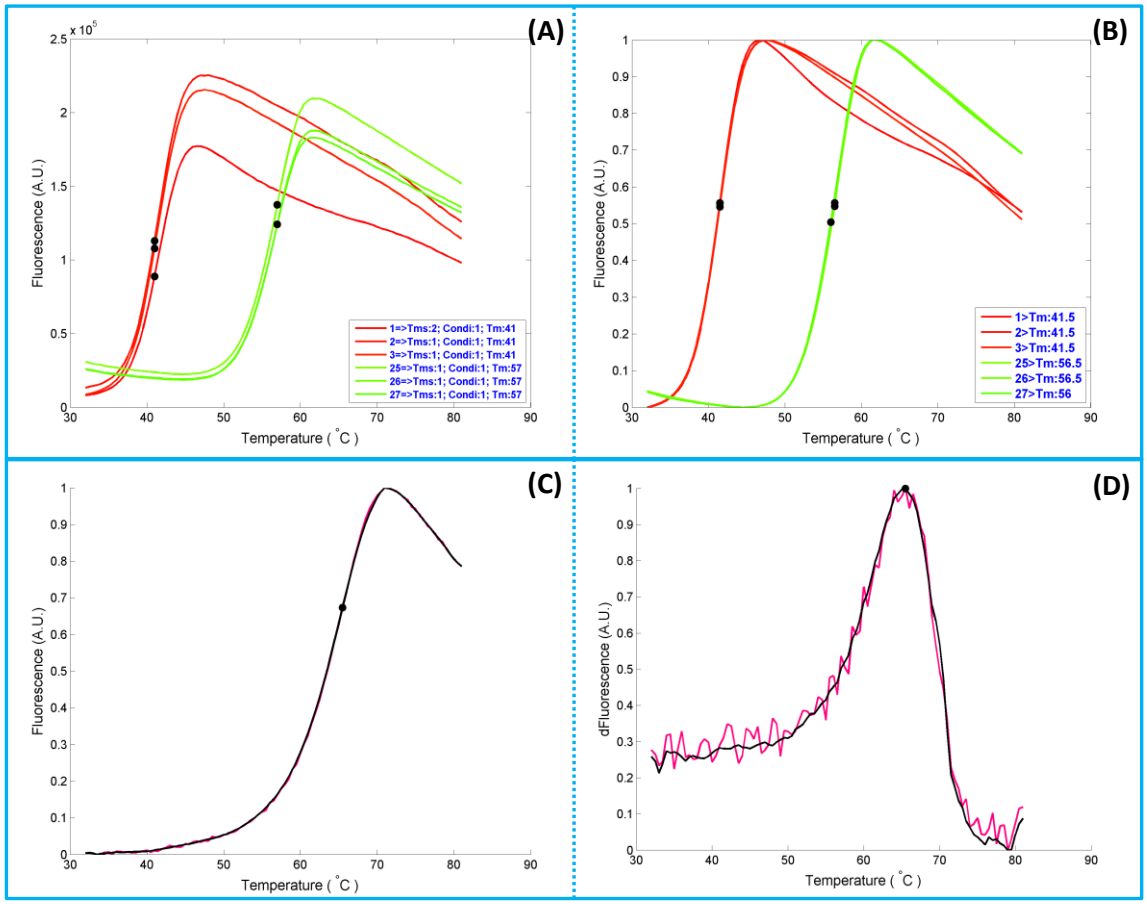


Figure 2 Normalise and smooth the melting curves. Normalising the melting curves changes the display range of measured fluorescence intensity from 0 to 1, but the shape of the melting curve is unaltered. Smoothing the melting curves reduces the noise level, but also does not alter the shape of the curves. (A): Original melting curves of His-FGF10 (red) and His-FGF10 stabilised with 1.25 µM heparin (green). (B): Normalised melting curves of the same samples in (A). (C): Original melting curve (pink) and smoothed melting curve (black) of His-FGF7 stabilised with 50 µM heparin. (D): First derivative curves of melting curves in (C).

Equation 1: $$MC = \frac{MC - \min(MC)}{\max(MC) - \min(MC)}$$

Equation 2: $$FD = \frac{MC(n+1) - MC(n-1)}{2 * \Delta t}$$

MC: Melting curve values; min(*MC*): minimum value for the melting curve; max(*MC*): maximum value for the melting curve; MC(n+1) is the $(n+1)^{th}$ value in the melting curve; Δt is the difference in temperature of two neighbouring measurement points.

3.3 Calculate melting temperature – (1) First derivative curve method: The melting temperature is generally defined as the temperature at the maximum of the first derivative curve (Fig. 3 A). (2) Melting curve method: The melting temperature can also be determined by calculating the half maximal denaturation temperature (**DT_{50}**), which is temperature when the fluorescence intensity reaches 50% of highest fluorescence intensity (Fig. 3 B). The DT_{50} is a slightly higher melting temperature, because it calculates the first temperature after the ideal DT_{50}. So, the difference of the ideal DT_{50} and acquired DT_{50} is less than the temperature increase for each measuring step.

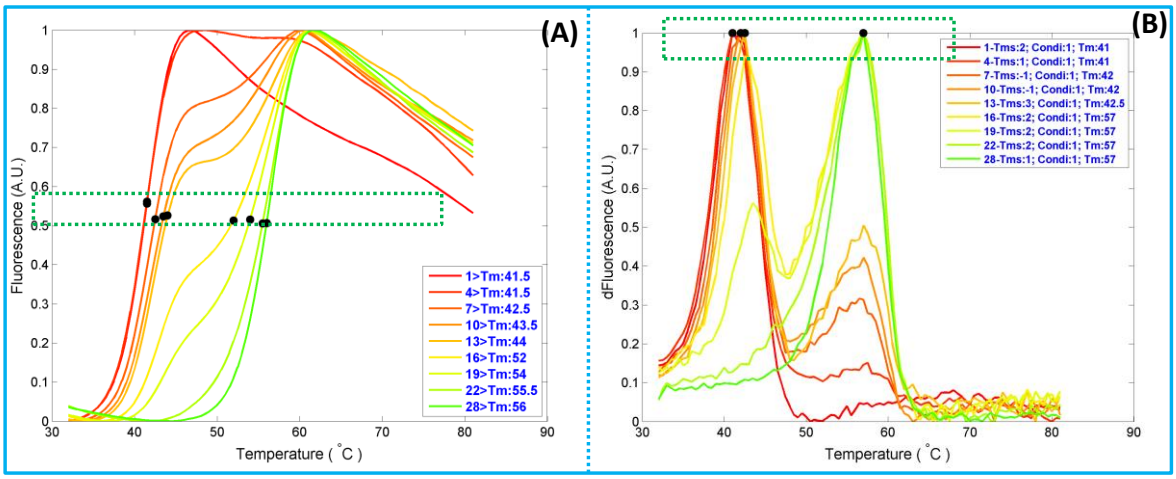


Figure 3 Extraction of melting temperatures. The stabilisation of His-FGF10 in the presence of a series of heparin ligands (concentrations are listed in Supplementary) was measured. The melting temperatures may be calculated by (A) the temperatures corresponding to half maximal fluorescence intensities and by (B) the temperatures corresponding to the maxima of the first derivative curves. The black points corresponding to melting temperatures are highlighted in the green dashed line boxes.

3.4 Auto-screen a melting curve – The melting curves are screened by the default parameters when the data are input into the program. If the melting curve decreases

from the maximal value to the minimal value as the temperature increases (Fig. 4 A, pink line), the protein corresponding to the melting curve is considered to be a denatured or unfolded protein. The *findpeak* function is used to find the peaks of the first derivative curves (Fig. 4 B). Only when the width and height of a peak are over 10 % that of the highest peak in the assay, is the peak recognised as a second melting temperature (Fig. 4 B, orange line). However, if the number of recognised peaks is >=4, the corresponding protein is also thought to be a denatured protein (no example shown).

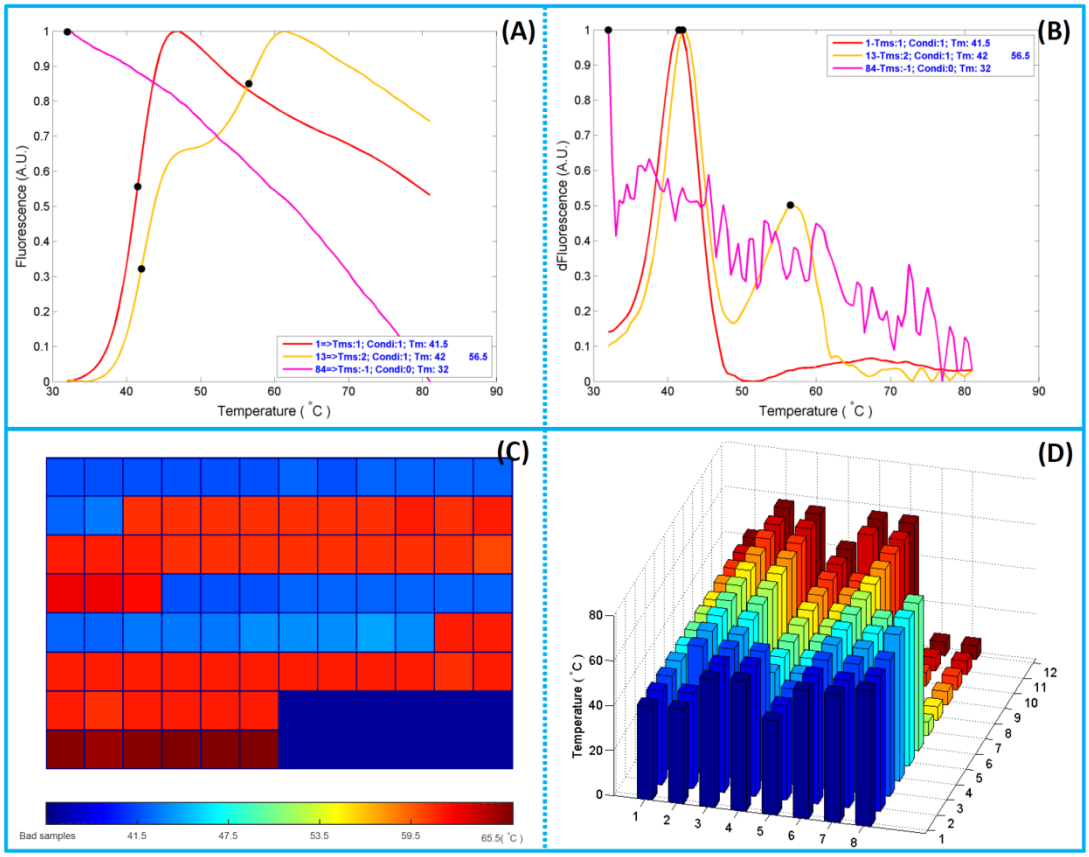


Figure 4 Auto-screening the melting curves and preview of melting temperatures in 96 wells. Three different types of melting curves (A) can be screened and recognised by the program. The screened results and the melting temperature information are automatically presented in two figures (C and D). (A): Melting curves of His-FGF10 (red line), His-FGF10 stabilised with 0.625 μM heparin (yellow line) and His-FGF10 (pink line) denatured by 8 M urea. (B): First derivative curves calculated from melting curves in (A). (C): 2D map view of melting temperatures of the proteins in 96 wells. The dark blue wells contain unfolded proteins or unstable proteins (e.g., proteins in an inappropriate buffer). The

colours of the melting temperatures of folded proteins are distributed in the colour bar, which is from the lowest melting temperature to the highest temperature. (D): 3D view of melting temperatures. The melting temperature of each well is plotted with a bar in the 3D graph with same colour for each column.

3.5 Preview of the protein stabilities in the 96 wells – This function shows the wells containing unfolded proteins and to present the difference of their melting temperatures as an image map. The wells coloured with a low value colour (at about beginning of the colour bar, Fig. 4 C) are supposed to contain the unfolded proteins (Fig. 4 C). The melting temperatures are presented as different coloured squares from the lowest melting temperature the highest melting temperature (colour bar). A 3D bar plot containing the melting temperatures of the 96 wells is also prepared, to give a view of difference of proteins' thermal stabilities (Fig. 4 D).

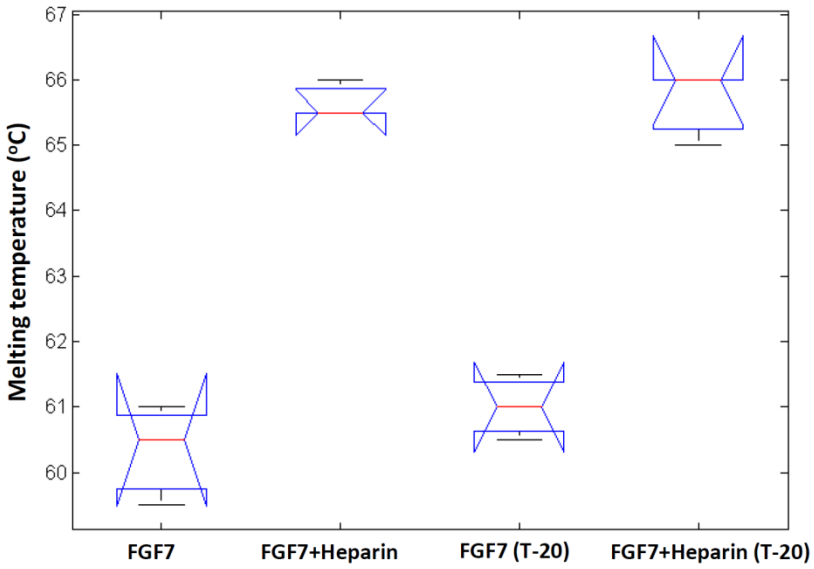


Figure 5 Statistical analysis of different samples with ANOVA. The melting temperatures of different samples can be compared with ANOVA. The melting temperatures of His-FGF7 in PBS, His-FGF7 stabilised with 50 µM heparin in PBS, His-FGF7 in PBS containing 0.5% (v/v) Tween-20 (T-20) and His-FGF7 stabilised with 50 µM heparin in PBS containing 0.5% (v/v) Tween-20 (T-20).

3.6 Comparison of melting temperatures of different samples – The melting temperatures of different samples can be compared by analysis of variance (ANOVA), as is described in Supplementary (User instruction). This will provide an appropriate statistical analysis (Fig. 5). The FGF7 example indicates that the addition

of 0.5% (v/v) Tween-20 did not change the melting temperature of FGF7 and heparin-stabilised FGF7, and the melting temperatures are the same as found previously (Xu et al., 2012).

4 Discussion

This program provides for analysis of large amounts of data, which can be easily transferred into analogous assays, such as spectral data acquired on 96-well plates. The program makes it more efficient to rapidly view and inspect the analysed data and produce a report graph. Further development can be introduced by the user or by contacting the authors.

5 Supplementary (Program details)

5.1 File Table:

(1) Code files (Folder: SimpleDSFviewer MATLAB code**):**

SimpleDSFviewer.m (SimpleDSFviewer.fig) is the main script, which displays the user interface and contains the buttons and parameter input.

Normalisemc.m is connected to the user interface script to normalise the melting curves by using the parameters loaded into the user interface.

Smomcurve.m is used to smooth the melting curves if it is required.

Peaknumber.m is used to screen the melting curves, which automatically classifies the curves. The different curves shown in Fig. 4 A can be recognised automatically, and the parameters for classifying the curves can be changed manually.

mapTm.m is to generate the preview map (Fig. 4 C) and 3 D bar plot graph (Fig. 4 D).

(2) Runtime app files (Folders: SimpleDSFviewer Mac/Win32/Win64 runtime**):**

SimpleDSFviewer.exe is the software package for running the program. Choose the appropriate version, depending on the computer (Windows 32 or Windows 64 or Mac), and double click this file to run the program.

readme.txt contains the details which introduce the requirement of installation of MATLAB runtime.

Other files are icon pictures and supporting files.

(3) Example data ('example raw.xlsx' and 'example sorted.xlsx'**).**

All the above files for the DSF data analysis program can be freely downloaded from GitHub: https://github.com/hscsun/SimpleDSFviewer.git.

5.2 User instruction:

5.2.1 Start the programme: Open SimpleDSFviewer.m in MATLAB R2015a (containing 'Signal Processing Toolbox', 'Statistics and Machine Learning Toolbox' and 'Curve Fitting Toolbox') and run the script. Alternatively, install the desired MATLAB Runtime 2015 (download from webpage: http://uk.mathworks.com/products/compiler/mcr/) and open the compiled app (SimpleDSFviewer), if MATLAB software is not installed. The interface (Fig. S1) will be loaded and ready to use. A Windows computer is preferred, since the analysed data cannot be written into an excel file for 'Save Data' function on a Mac computer.

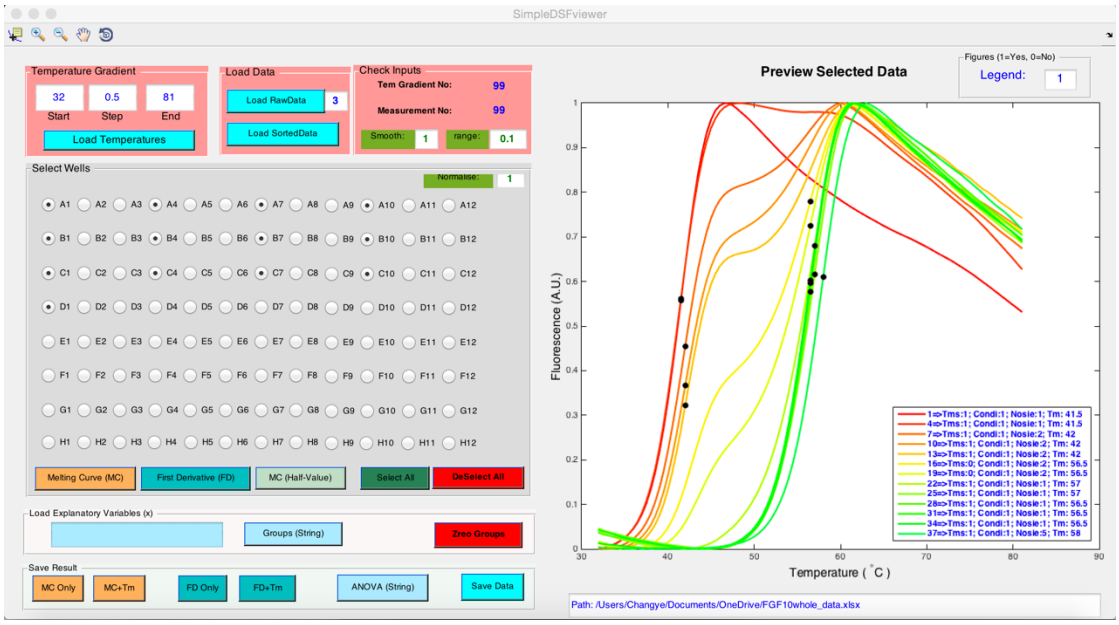


Figure S1 User interface of SimpleDSFviewer. This figure shows the user interface of SimpleDSFviewer, in which the user-can analyse and view the collected DSF data. The top-left three panels are used to load the parameters and collected DSF data. The 96 well radio buttons and the data view buttons are distributed on the bottom left. The right side contains a graphical view box for displaying the selected data and the path of input data in a text box.

5.2.2 Load parameters: The temperatures, including start temperature, temperature step and end temperature are required, if a raw data file (Fig. S2 A) is to be loaded. In the software, '1' signifies Yes and 0 signifies No for '*Smooth*', 'Normalise' and 'Legend' parameters. The *Smooth* and *Normalise* parameters should be kept as '1' as the data are loaded, but the range (*Smooth Range*) can be changed to be increased or decreased. The '*Normalise*' and '*Smooth*' functions can be turned off when the original melting curves are desired.

The default values are loaded automatically and are generally ideal for analysis.

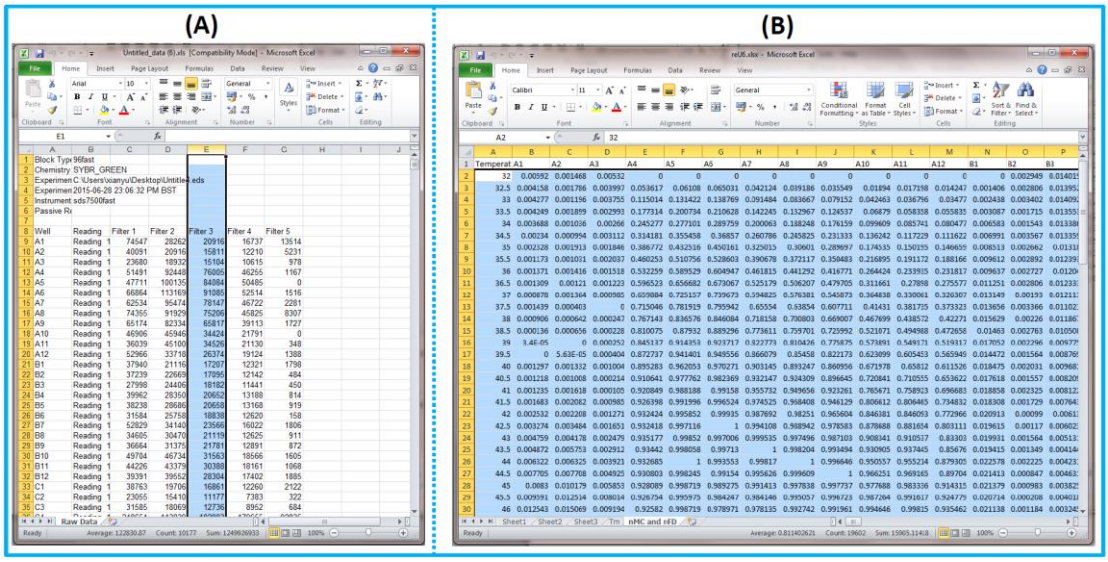


Figure S2 Templates of input data for SimpleDSFviewer. Two formats of input data can be loaded into the program. (A): Raw data contains all the fluorescence reading from wells 1 to 96 and from the lowest temperature to the highest temperature in one column. The number of desired column needs to be specified. (B): Sorted data contains the melting temperature in the first column and the 96 wells' fluorescence in 96 columns.

5.2.3 Load experimental result: Two formats of data can be loaded, as shown in Fig. S2. The first format is the raw data, which contains all the data in a single column and the column that contains the desired data should be specified (Text box next to '*Load RawData'* pushing button, Fig. S1). Column 3 was selected in the example (Figs. S1 and S2 A). The second format is sorted data (Fig. S2 B), which contains 96 columns corresponding to 96 wells and fluorescence intensities of each temperature in each row. The loading data should be stored in the first sheet of the Excel file.

5.2.4 Functions: Multiple wells can be selected by clicking the corresponding radio buttons for data view. The melting curves and first derivative curves can be previewed in the graph box (Fig. S1) by pushing the buttons (containing MC or FD)

in the *Select Wells* panel or be viewed and saved by pushing the buttons in the *Save Result* panel. '+T_m' means the melting temperature is marked on the melting curve. MC (Half-Value) shows the melting curves with half denaturation point (main paper 3.3). The legend can be added by changing the legend parameter to '1'.

Different samples can be compared by ANOVA: Repeating samples are selected by clicking the radio buttons and a name should be input the text box in the '*Load Explanatory Values (X)*' panel. Then, the sample is added by pushing '*Group (String)*' button. The selected wells should be deselected and more samples can be added in the same way. The ANOVA (string) button will give the statistical analysis and a box-plot graph for the samples under comparison.

5.2.5 Save function: The sorted (Sheet1) and normalised melting curves (Sheet: nMC and nFD) and the first derivative curves and the melting temperatures (Sheet: T_m) calculated by the three different methods (Fig. 1) may be saved as an Excel file by pushing '*Save Data'* button.

5.2.6 Errors: When the data are input into the programme, a running status bar will be displayed on the screen and it will be turned off after the auto-analysis. However, if the running cannot be completed (bar status), the input data should be checked. For example, if the gradient numbers of temperature are longer than the fluorescence measuring numbers for each well, an error will be generated and running will not complete. Missing values in the loaded data also cause the same problem.

5.2.7 Example data: Protein and the heparin (Hep) ligand concentrations are listed in the following table. Every three consecutive wells contain three same samples to ensure the experimental results.

FGF10 PBS	FGF10 Hep 0.2 µM	FGF10 Hep 0.4 µM	FGF10 Hep 0.5 µM
FGF10 Hep 0.625 µM	FGF10 Hep 0.75 µM	FGF10 Hep 0.875 µM	FGF10 Hep 1 µM
FGF10 Hep 1.25 µM	FGF10 Hep 2.5 µM	FGF10 Hep 5 µM	FGF10 Hep 50 µM
FGF10 Hep 500 µM	FGF10 PBS	FGF10 Hep 0.2 µM	FGF10 Hep 0.4 µM
FGF10 Hep 0.5 µM	FGF10 Hep 0.625 µM	FGF10 Hep 0.75 µM	FGF10 Hep 0.875 µM
FGF10 Hep 1 µM	FGF10 Hep 1.25 µM	FGF10 Hep 2.5 µM	FGF10 Hep 5 µM
FGF10 Hep 50 µM	FGF10 Hep 500 µM	FGF10 6 M urea	FGF10 8 M urea
FGF7 PBS	FGF7 Hep 50 µM	FGF7 0.5% Tween-20 (v/v) PBS	FGF7 0.5% Tween-20 (v/v) Hep 50 µM

4.3 Binding selectivity of FGFs to heparin derivatives

4.3.1 Proteins and ligands

FGF10 and FGF20 (5 μM) were used to study their interactions with the following a library of chemically-modified heparins, oligosaccharides and GAGs.

(a): Porcine intestinal mucosal heparin (Sigma-Aldrich, Dorset, UK, concentrations are listed in the figure legend of Fig. 4.1 A). As described in Section 1.4, heparin is a highly sulfated HS, which is nonetheless very polydisperse, both in terms of polymer length and sulfation. The average molecular weight of the heparin is 17.5 kDa, with a range that is generally between 5 kDa and 30 kDa, though some smaller and larger material is also present.

(b): Heparin oligosaccharides: dp2, dp4, dp6, dp8, dp10 and dp12 (10 μM, Iduron, Manchester, UK).

(c): Heparin derivatives shown in Table 4.2 (10 mM, gift from Dr. Ed Yates, University of Liverpool).

(d): GAGs: HS (10 μM, Sigma-Aldrich), CS (10 μM, Iduron), DS (10 μM, Iduron) and HA (10 μM, Sigma-Aldrich). The structures of the constituent disaccharides of these GAG chains are described in Figure 4.1.

4.3.2 Experimental setup and data analysis

FGF10, FGF20 and the GAG ligands were mixed with SYPRO dye by following table 4.1. The measurement and data analysis were carried out by following Section 4.2.

Table 4.1 DSF reaction components for three wells:

Components	FGF	GAG ligand (10X)	SYPRO orange (100X)	PBS	Total volume
Volume / μL	a[1]	3.5	3.5	(28-a)	35

a[1]: volume of the protein for specified concentration.

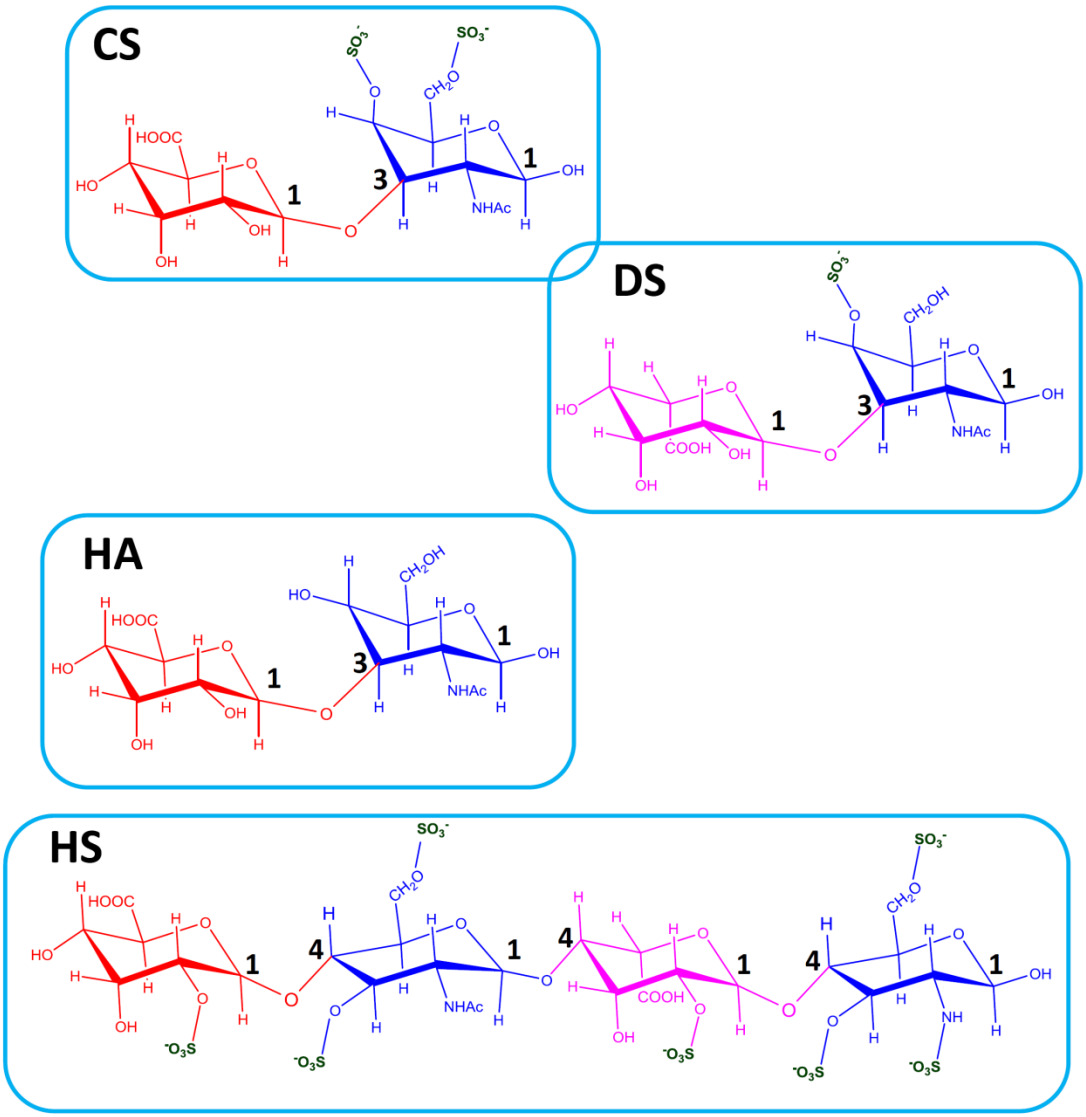


Figure 4.1 Structures of CS, DS, HA and HS (Goel and Gowda, 2011, Yung and Chan, 2007). Red saccharide: glucuronic acid; pink saccharide: iduronic acid; blue saccharide: N-acetyl-D-galactosamine (in CS and DS), N-acetyl-D-glucosamine (in HA and HS) or N-sulfated D-glucosamine (in HS). In CS and HA, a β(1–3) glycosidic bond links glucuronic acid to galactosamine or glucosamine, while in DS at least one glucuronic acid is epimerised to iduronic acid. In all these three GAGs, β(1–4) glycosidic bonds connect galactosamine (or glucosamine) to glucuronic acid (or iduronic acid). In contrast, the saccharide residues are linked by β(1–4) or α(1–4)

glycosidic bonds in HS. The possible sulfation positions are all labelled in the structures (green).

Table 4.2 Chemically modified heparin structures (Xu et al., 2012):

The letter I stands for iduronate (IdoA), and A stands for the amino sugar glucosamine (GlcN). Numbers refer to the carbon of the monosaccharide.

Name	Predominant repeat	IdoA-2	GlcN-6	GlcN	IdoA-3	GlcN-3
Heparin	$I^{2S}A^{6S}_{NS}$	$-SO_3^-$	$-SO_3^-$	$-SO_3^-$	-OH	-OH
D2	$I^{2S}A^{6S}_{NAc}$	$-SO_3^-$	$-SO_3^-$	-Ac	-OH	-OH
D3	$I^{2OH}A^{6S}_{NS}$	-OH	$-SO_3^-$	$-SO_3^-$	-OH	-OH
D4	$I^{2S}A^{6OH}_{NS}$	$-SO_3^-$	-OH	$-SO_3^-$	-OH	-OH
D5	$I^{2OH}A^{6S}_{NAc}$	-OH	$-SO_3^-$	-Ac	-OH	-OH
D6	$I^{2S}A^{6OH}_{NAc}$	$-SO_3^-$	-OH	-Ac	-OH	-OH
D7	$I^{2OH}A^{6OH}_{NS}$	-OH	-OH	$-SO_3^-$	-OH	-OH
D8	$I^{2OH}A^{6OH}_{NAc}$	-OH	-OH	-Ac	-OH	-OH
D9	$I^{2S,3S}A^{6S}_{3S,NS}$	$-SO_3^-$	$-SO_3^-$	$-SO_3^-$	$-SO_3^-$	$-SO_3^-$

The difference of melting temperatures of the FGF in PBS buffer and in the presence of ligand (heparin and heparin derivatives) was used to show the relative stabilising effect (Eq1) (Uniewicz et al., 2010).

$$\text{Eq1: Relative stabilising effect} = \frac{\text{Tm(FGF + ligand)} - \text{Tm(FGF + heparin)}}{\text{Tm(FGF in PBS)} - \text{Tm(FGF + heparin)}}$$

4.3.3 Results and discussion

Thermal stability of FGF10 and FGF10-sugar complexes: The melting curve of FGF10 shows the fluorescence intensity was increased from ~40ºC and reached a maximum at about 46ºC (Fig. 4.2 A). As heparin was added, FGF10 was stabilised by binding to heparin, which shifted the melting curve of FGF10 to higher temperatures (Fig. 4.2 A). The first derivative curve reflects the melting temperature of each protein (Section 4.2). Interestingly, these first derivative curves show that only small amounts of FGF10 were stabilised (the second peak), when 0.4 and 0.5 µM heparin was added. This suggests that FGF10 has a slow disassociation rate from heparin, so that there is no averaging of the bound and unbound FGF10 on the time scale (~5 min for the denaturation) of the measurement. Thus, at subsaturating concentrations of heparin, FGF10 is either bound or not bound to heparin, and the melting of both species is seen simultaneously. In contrast, in previous work with other FGFs (FGF1, FGF2, FGF7, FGF9 and FGF18), at these concentrations of heparin, just one intermediate melting temperature was observed (Uniewicz et al., 2010, Xu et al., 2012), since the proteins would exchange on the polysaccharide over the time of the measurement.

The melting temperature of FGF10 in PBS buffer is around 42°C, which is below that of many other reported FGFs (FGF1, FGF2, FGF7, FGF9 and FGF18) (Xu et al., 2012, Uniewicz et al., 2010). The binding of FGF10 to heparin increases its thermal stability by ~15°C (Fig. 4.2: A-C), which is considerably more than seen with FGF7 (~7°C), but less than the ≥20°C shift seen with FGF1 and FGF2 (Uniewicz et al., 2010, Xu et al., 2012, Xu, 2012). Maximum stabilisation of FGF10 (~5 µM) was attained with 1 µM heparin, suggesting that more than one FGF10 could bind to a

single heparin chain. FGF10 did not interact with a dp2, with binding apparent from dp4 and a heparin-derived dp8 was as effective as full length heparin (Fig. 4.2 E). The melting curves of binding of FGF10 to heparin oligosaccharides show that FGF10 was more stabilised as increasing polymer length from dp4 to dp10 (Fig. 4.3). Since there was only one peak in each first derivative curve, all of the FGF10 in the reaction was stabilised by oligosaccharides (Fig. 4.3 B). Therefore, the gradual change of melting temperatures of FGF10-oligosaccharide was due to the partial stabilisation of FGF10 by short heparin and a dp8 or dp10 is preferred for binding. This is consistent with heparin having an average molecular weight of 17.5 kDa, which corresponds to around dp60 and the observed stoichiometry of five FGF10 per heparin (Fig.4.2C).

The relative stabilising effect of FGF10 by heparin derivatives was measured to identify its preference for particular sulfation patterns. FGF10 did not discriminate between any of the doubly sulfated heparins (D2-D4) (Fig. 4.2 D). It bound less well to singly sulfated heparins and showed a modest preference for 6-O sulfated heparin over 2-O sulfated and perhaps N sulfated heparin, and failed to bind to completely desulfated heparin (Fig. 4.2 D). FGF10 bound persulfated heparin more strongly than heparin, and its interaction with HS was similar to that seen with singly 6-O sulfated heparin. Interestingly, FGF10 bound, though weakly, to CS and more strongly to DS, showing that there is a wider range of GAG species it can recognise (Fig. 4.2 D).

In terms of cation coordinated heparins, FGF10 had a slight preference for divalent cation forms of the polysaccharide over the monovalent cation forms (Fig. 4.2 F). Compared to FGF7 (Table 4.3, Section 4.4), FGF10 is less sensitive to the position and number of sulfate groups in the polysaccharide, though it binds persulfated

heparin more strongly and exhibits a preference for divalent cation forms of heparin (Xu et al., 2012). The latter difference may be important for tuning responses in situations where both FGF7 and FGF10 are present, since previous studies have shown that cation coordination by heparin can alter FGF-heparin interactions and their biological activities (Rudd et al., 2007, Kan et al., 1996). Both FGF7 and FGF10 can bind to CS and DS, which makes them relatively unique amongst the FGFs. Taken together, the data indicate that FGF10 binds to sulfated structures in GAGs in a similar way to FGF7, but that FGF10 is more stabilised by heparin than FGF7.

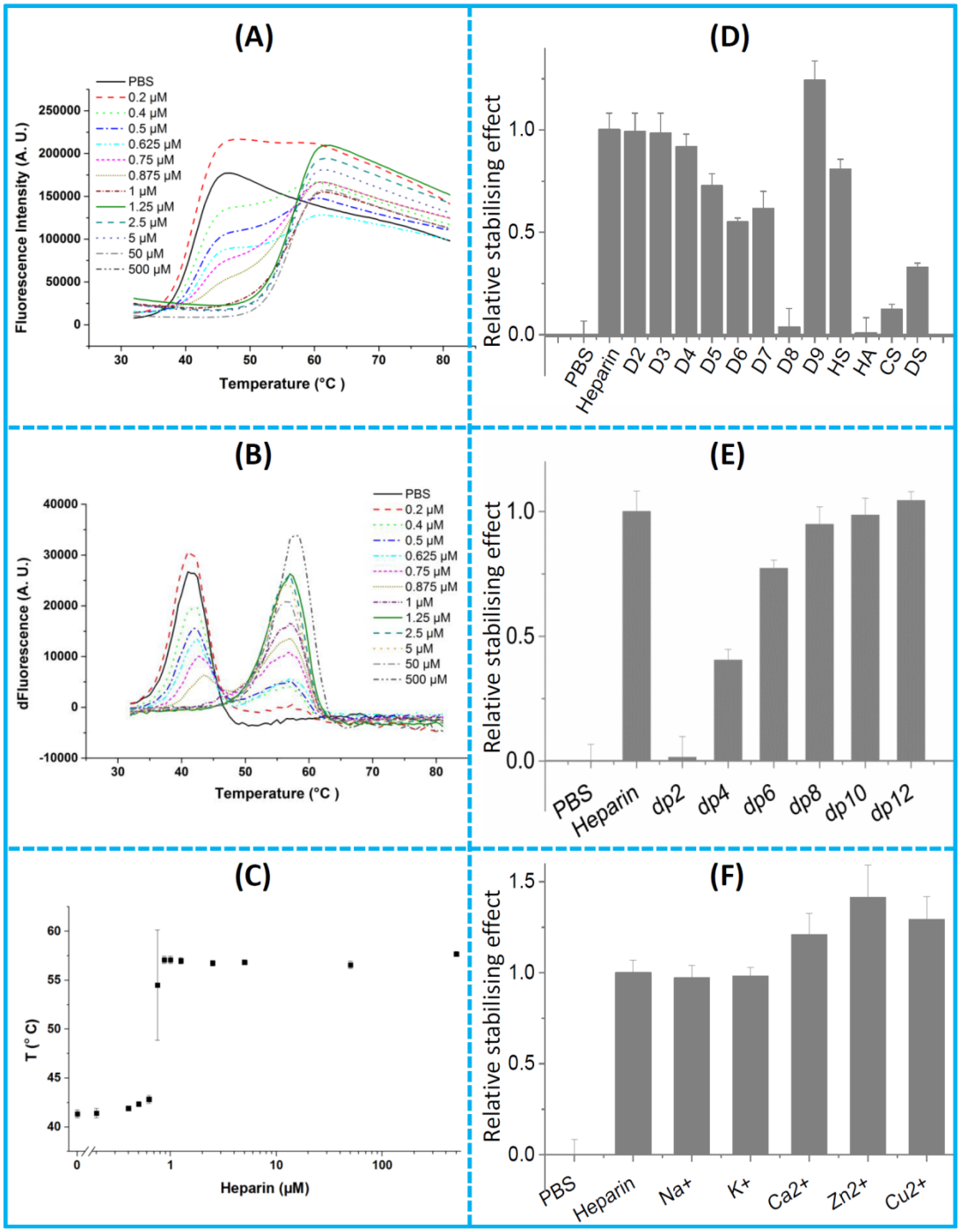


Figure 4.2 Stabilisation of FGF10 (5 μM) by modified heparins and model GAGs. (A): Melting curves of FGF10 with different concentrations of heparin. **(B):** First derivative of the melting curves in (A). **(C):** Melting temperatures of FGF10 with different concentrations of heparin. **(D):** Relative stabilising effect of FGF10 by different heparin derivatives, HS, HA, CS and DS (all 10 μM). **(E):** Relative

stabilising effect of FGF10 by heparin-derived oligosaccharides dp2 to dp12 (all 10 µM). **(F):** Relative stabilising effect of FGF10 by heparin (10 µM) coordinated by different cations (Na^+, K^+, Ca^{2+}, Zn^{2+} and Cu^{2+}). Relative stabilising effect was calculated by following Eq1 (Section 4.3.2).

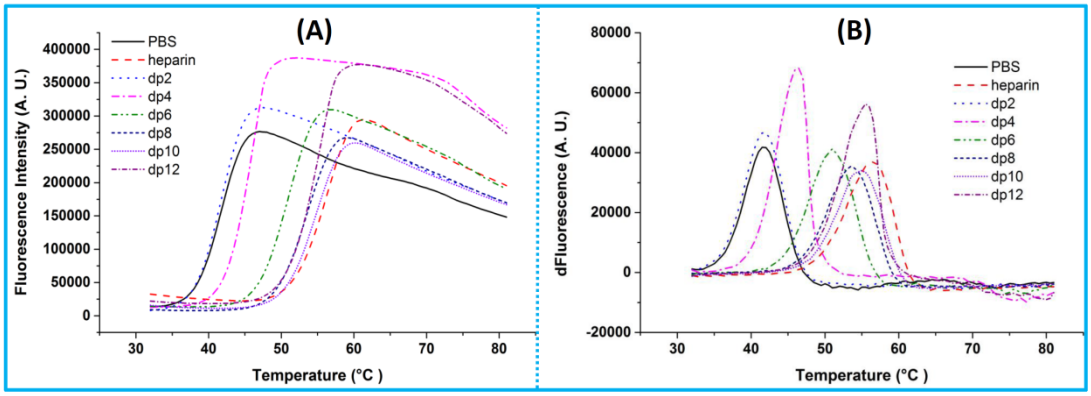


Figure 4.3 Binding of FGF10 (5 µM) to heparin oligosaccharides (10 µM). (A): Melting curve of FGF10 in PBS and in PBS containing 10 µM heparin and heparin-derived oligosaccharides of different lengths (dp2-dp12 and full length heparin); **(B):** First derivative of the melting curves in (A).

Thermal stability of FGF20 and FGF20-sugar complex: The melting curve of FGF20 was increased from about 45°C and the addition of heparin shifted the melting curve of FGF20 to higher temperatures (Fig 4.4 A). The first derivative curve indicates that the melting temperature of FGF20 in PBS buffer is about 52°C (Figs 4.4 B-C), which is consistent with previous studies (Fan et al., 2007, Maity et al., 2009). FGF20 was strongly stabilised by binding to heparin (Fig. 4.4: A-C), but in contrast to FGF10, FGF20 (~5 µM) did not appear to reach full stabilisation, and even at 10 µM heparin the stabilisation was not complete. The average length of a heparin chain (>25 nm for dp60) is much greater than that required for binding (two

aligned canonical HBSs in a FGF20 dimer would be around 7 nm). Because heparin is polydisperse, this suggests that the requirement for a high concentration of heparin may be due to FGF20 preferentially binding a rare structure in the polysaccharide; the effective concentration of heparin chains with such structures would then be lower than that of total heparin. Since FGF20 is a dimer, such structural requirements may reflect conformations available to the heparin chain, which are strongly affected by substitution by sulfate groups (Rudd et al., 2007).

In terms of the length of heparin oligosaccharide that will bind to FGF20, the pattern is quite distinct. The melting curves of FGF20-oligosaccharide complexes indicate that the short modified heparin (dp2-dp8) did not stabilise FGF20 (Figs 4.4E and 4.5). Only a small stabilisation was observed with heparin-derived dp10 and about 40% of the relative stabilising effect of heparin was achieved by a dp12 oligosaccharide (Fig. 4.4 E). These data indicate that FGF20 either requires longer structures to bind (around 7 nm to 8 nm, equivalent to dp16) or that the structures it binds are rare in heparin and/or tend to be destroyed when heparin is partially digested with heparinase I (cleaves at IdoA2S) to make the oligosaccharides. Since heparin itself, though sufficiently large, cannot fully stabilise FGF20, it is more likely that FGF20's binding structures are rare in this polysaccharide.

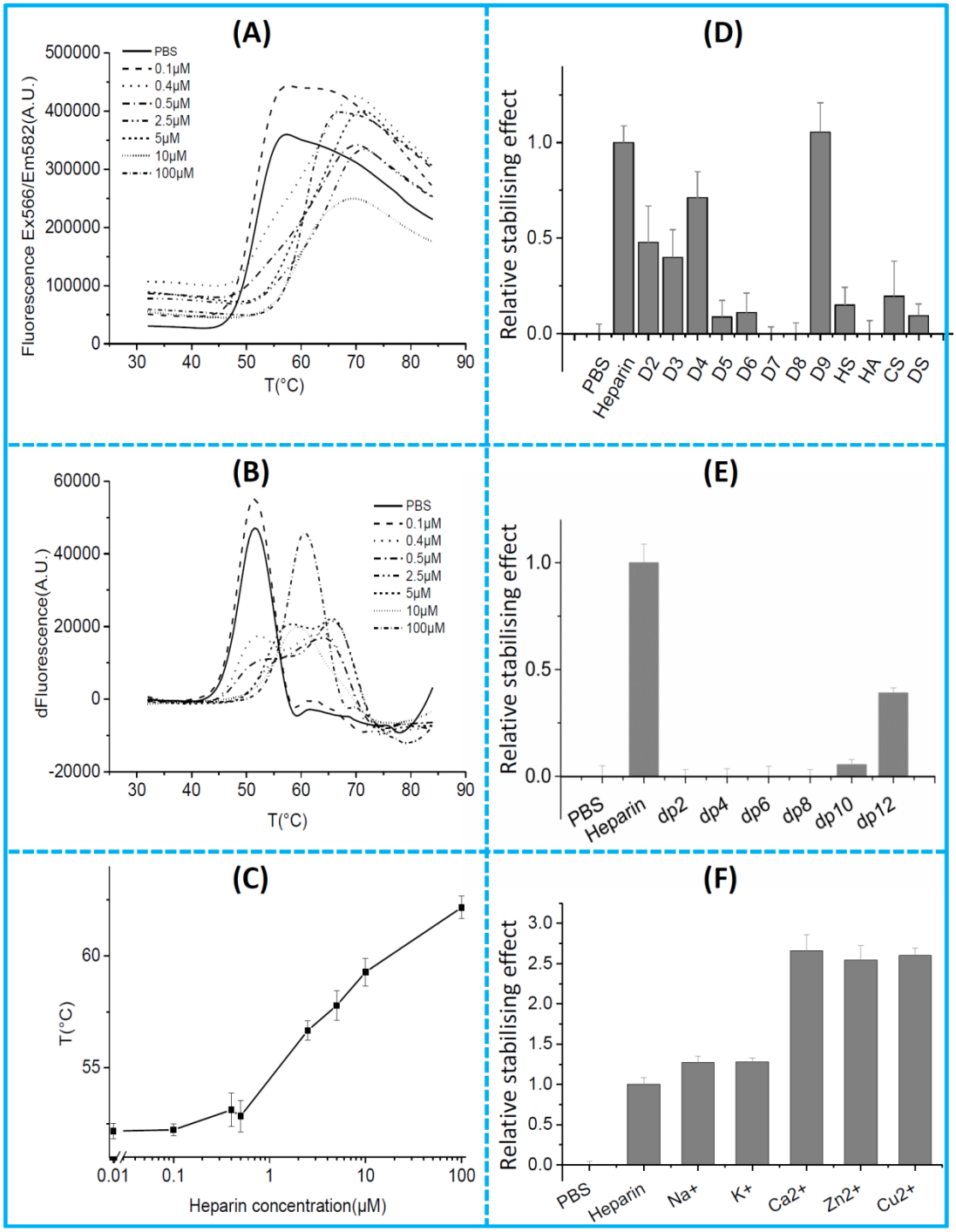


Figure 4.4 Stabilisation of FGF20 (5 μM) by modified heparins and model GAGs. (A): Melting curves of FGF20 with different concentrations of heparin. **(B):** First derivative of the melting curves in (A). **(C):** Melting temperatures of FGF20 with different concentrations of heparin. **(D):** Relative stabilising effect of FGF20 by different heparin derivatives, HS, HA, CS and DS (all 10 μM). **(E):** Relative

stabilising effect of FGF20 by heparin-derived oligosaccharides dp2 to dp12 (all 10 µM). **(F):** Relative stabilising effect of FGF20 by heparin (10 µM) coordinated by different cations (Na^+, K^+, Ca^{2+}, Zn^{2+} and Cu^{2+}). Relative stabilising effect was calculated by following Eq1 (Section 4.3.2).

FGF20 binding was reduced by around half when one sulfate was removed from heparin (D2-D4) and was marginal or undetectable with heparins possessing just one sulfate (D5-D7; Fig. 4.4 D). There was a slight preference of FGF20 for 2-O- and N-sulfate groups over 6-O-sulfate group. Binding to HS was weak and similar to the interactions with the singly sulfated D5 and D6, and the two model GAGs, CS and DS (Fig. 4.4 D). FGF20 has a marked preference for divalent cation coordinated heparins, since these stabilised the protein more than twice as effectively as heparin (Fig. 4.4 F). Although the cation coordinated status of HS *in vivo* is unknown, it would seem that divalent cations might have some important roles in the regulation of the biological activities of FGF20.

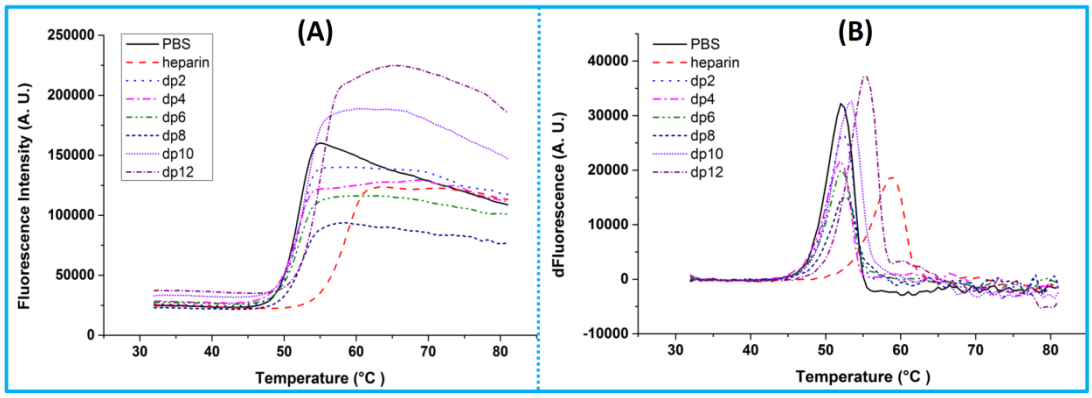


Figure 4.5 Binding of FGF20 (5 µM) to heparin oligosaccharides (10 µM). (A): Melting curve of FGF20 in PBS and in PBS containing 10 µM heparin and heparin-

derived oligosaccharides of different lengths (dp2-dp12 and full length heparin); **(B):** First derivative of the melting curves in (A).

4.4 Conclusion

In this chapter, DSF was used to study protein stability and protein-ligand interactions, and a computer program 'SimpleDSFviewer' was developed to analyse DSF data. This provides far faster analysis than the conventional manual step-by-step analysis.

FGF10 is a member of the FGF7 subfamily. The data on the interactions of FGF10 with heparin and heparin derivatives indicate that it shares most of its preferences with FGF7, but, like FGF7, is distinct from FGFs in other subfamilies that have been analysed so far (Table 4.3) (Xu et al., 2012, Uniewicz et al., 2010). The main difference is heparin can better stabilise FGF10 (~15^{o}C) than FGF7 (~7^{o}C). FGF20 is in the FGF9 subfamily and there are strong similarities in its preferences for binding particular structures in the model GAGs used here (Table 4.3). The main differences between FGF9 and FGF20 are that FGF20 is more extreme in terms of its requirement for structures longer than dp10, the weakness (relative to heparin) of its interactions with singly sulfated heparins and its very strong preference for divalent cation coordinated polysaccharide. However, similarly, for FGF9 dp10 and dp12 can only stabilise the protein to 50% of the level observed with heparin and the stabilisation effect can be significantly decreased by removing any sulfation group in heparin (Xu et al., 2012). Therefore, FGF20 is more similar in its binding preferences to FGF9 than to FGFs in other subfamilies that have been characterised (Uniewicz et al., 2010, Xu et al., 2012). Previous work provides data on two

members of one other FGF subfamily, FGF2 and FGF1 (Table 4.3). While these FGFs have different sulfation preferences for binding, in that FGF1 is more promiscuous and binds to all singly desulfated heparins, whereas FGF2 requires N- and 2-O sulfate groups, they are both strongly stabilised by heparin and have similar relative stabilising effects by dp4-dp10.

Table 4.3 Summary of binding preference of FGFs to heparin, heparin derivatives and heparin oligosaccharides (Uniewicz et al., 2010, Xu et al., 2012). Red coloured FGFs are from previous publications and green FGFs are from this PhD project.

Subfamily	Name	Heparin effect	Sulfation preference	dp length
FGF1	FGF1	> 22°C	NS ≈ 2S ≈ 6S	dp2 ~ dp10
	FGF2	> 22°C	NS, 2S > 6S	dp4 ~ dp10
FGF7	FGF7	~ 7°C	NS ≈ 2S ≈ 6S	dp4 ~ dp8
	FGF10	~ 15°C	NS ≈ 2S ≈ 6S	dp4 ~ dp8
FGF8	FGF18	> 15°C	NS ≈ 6S > 2S	dp6 ~ dp10
FGF9	FGF9	~ 20°C	NS ≈ 6S > 2S	dp4 ~ dp10
	FGF20	~ 12°C	6S ≈ 2S > NS	dp10 ~ dp12

Taken together the present dataset on FGF10 and FGF20 and the previous findings supports the overall hypothesis that HS binding selectivity is more conserved in FGF subfamilies, and so is likely to have been subjected to the same natural selection pressures that led to the diversification of the FGF family and the more complex body plan of vertebrates.

Chapter 5 Binding and diffusion of Halo-FGFs in Rama 27 fibroblast pericellular matrix

5.1 Introduction

Gradients of signalling proteins are important in development, growth and homeostasis. For example, in embryonic development, morphogen gradients that form in the ECM are important to achieve functions, such as cell specification and differentiation. Morphogens are expressed and secreted into the ECM where many of them can bind to HS. Many other signalling proteins share this property and they are part of the heparin/HS-binding proteins, of which there are 883 have been identified in the human proteome (Ori et al., 2008, Ori et al., 2011, Nunes, 2015). HS chains anchored to the core proteins of the proteoglycans are proposed to be an important regulator of all HS-binding proteins and of the movement of morphogens and other signalling proteins (Sections 1.2 and 1.8).

However, it is still unclear how and to what extent HS in the ECM can regulate the transport of these proteins. In particular, there is little evidence to indicate that HS has much specificity or selectivity in this function. The FGFs are a key model for addressing these questions, because they are the first recognised "heparin binding growth factors", so a considerable amount of work on the molecular specificity and the functional significance of their interactions with HS has been done. Together with the likelihood that the specificity of FGFs for binding structures in HS was subjected to the same natural selection pressures that led to the expansion of the FGF

family (Chapter 4), this means that FGFs are a powerful tool to determine the extent to which HS may be able to differently regulate FGF diffusion (Chapter 4).

A limited number of studies have measured the movement of FGFs through ECM to shed light on the formation of gradients of FGFs (Section 1.9). A study with FGF7, FGF10 and FGF10 mutants (exchanging arginine (193) with lysine or arginine (187) with valine in the heparin binding site) suggests that FGF10 binds to heparin much more strongly than FGF7 and the FGF10 mutants, so that it does not diffuse far or fast from heparin beads (Makarenkova et al., 2009). This has been interpreted to explain why FGF10 causes elongation of lacrimal and salivary gland ductal epithelium, whereas the weak binding FGF7 gives rise to branching (Makarenkova et al., 2009), though other work also points out that heparanase, which will release S-domains from HS, is critical to the ductal morphogenesis activity of FGF10 in salivary gland (Patel et al., 2007).

Previous work in the pericellular matrix of Rama 27 fibroblasts had suggested that the distribution of FGF2 binding sites in HS was not homogenous, suggestive of a higher-order organisation of this matrix. Moreover, measurement of the movement of the FGF2 in this pericellular matrix demonstrated that the FGF2 could undergo different types of movement, from confined motion to directed diffusion over many µm (Duchesne et al., 2012). What is not clear is whether these properties were specific to just one FGF (or indeed just one heparin/HS binding protein) or were more general. Thus, if ECM presented binding sites for different proteins on HS in a clustered and organised manner, this would allow the independent control of diffusion of different HS binding proteins. For this to occur, it is likely that the ECM would have to possess a supramolecular organisation. That is a structure due to

interactions of ECM components that extends beyond the length scale of its individual components.

In this Chapter, two FGFs from one subfamily (FGF1 and FGF2, FGF1 subfamily) and three (FGF6, FGF10 and FGF20) from other subfamilies (FGF4, FGF7 and FGF9), all with well characterised polysaccharide binding properties, were used to determine how they bind to ECM and diffuse. To ensure that the data reflected properties of the ECM and were not confounded by cell biochemistry, fixed Rama 27 fibroblasts were used.

5.2 Binding and diffusion of FGFs in fibroblast pericellular matrix (Paper)

Sun, C., Marcello, M., Li, Y., Mason, D., Lévy, R., Fernig, D. G. (2015). Selectivity in glycosaminoglycan-binding dictates the distribution and diffusion of fibroblast growth factors in the pericellular matrix. (Submitted to *Open Biology on 23rd December 2015*)

Contributions of authors:

Changye Sun: Produced Halo-FGFs and FGF1, prepared TMR labelled Halo-FGFs, acquired imaging data with confocal microscopy, analysed the imaging data, prepared figures and wrote the paper. **Marco Marcello:** Taught the confocal microscopy, supported ideas for imaging and edited the paper. **Yong Li:** Provided FGF2 and edited the paper. **David Mason:** Supported ideas for imaging data analysis

and edited the paper. **Raphaël Lévy:** Conceived study and edited the paper. **David G. Fernig:** Conceived study and edited the paper.

Selectivity in glycosaminoglycan-binding dictates the distribution and diffusion of fibroblast growth factors in the pericellular matrix

Changye Sun[1], Marco Marcello[2], Yong Li[1], David Mason[2], Raphaël Lévy[1] and David G. Fernig[1]

Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool, UK L69 7ZB

Centre for Cell Imaging, Institute of Integrative Biology, University of Liverpool, Liverpool, UK L69 7ZB

Abstract:

The range of biological outcomes generated by many signalling proteins in development and homeostasis is increased by their interactions with glycosaminoglycans, particularly heparan sulfate (HS). This interaction controls the localisation and movement of these signalling proteins, but whether such control depends on the specificity of the interactions is not known. We used five fibroblast growth factors with an N-terminal HaloTag for fluorescent labelling (Halo-FGFs), with well-characterised and distinct HS binding properties, and measured their binding and diffusion in pericellular matrix of fixed rat mammary 27 fibroblasts. Halo-FGF1, Halo-FGF2 and Halo-FGF6 bound to HS, whereas Halo-FGF10 also interacted with chondroitin sulfate/dermatan sulfate, and FGF20 did not bind detectably. The distribution of bound FGFs in pericellular matrix was not homogenous, and for FGF10 exhibited striking clusters. Fluorescence recovery after photobleaching showed that FGF2 and FGF6 diffused faster, whereas FGF1 diffused more slowly, and FGF10 was immobile. The results demonstrate that the specificity of the interactions of proteins with glycosaminoglycans controls their binding and diffusion. Moreover, extracellular matrix has long-range structures, since cells regulate independently the distribution of different protein sites in glycosaminoglycans.

Introduction:

Extracellular matrix has a central role in mediating communication between animal cells through mechanisms mediated by mechanical forces and soluble effectors. A large proportion of the soluble effectors, morphogens, growth factors, cytokines and chemokines that regulate animal development and homeostasis interact with glycosaminoglycans, particularly heparan sulfate, of the extracellular matrix (Ori et al., 2008, Xu and Esko, 2014). These interactions have been shown to exhibit varying degrees of specificity and selectivity at the tissue and at the molecular levels, and in a number of cases have been demonstrated to control the effectors' transport and intracellular signalling.

The glycosaminoglycans heparan sulfate, chondroitin sulfate and dermatan sulfate are linear, sulfated polysaccharides covalently attached to core proteins to form proteoglycans. These are either associated with the cell membrane and resident in pericellular matrix or secreted, so resident in extracellular matrix. The long chains of heparan sulfate (~25 to 100 disaccharide units) consist of repeats of a disaccharide: D-glucosamine β 1-4 glucuronic acid or its epimer iduronic acid. The mature chains have a distinct domain structure of sequential blocks of unmodified disaccharides of N-acetyl glucosamine β 1-4 glucuronic acid, transition domains where N-acetyl glucosamine-containing disaccharides alternate with N-sulfated ones, and sulfated domains, where every glucosamine is N-sulfated and may also be O-sulfated on C_3 and C_6, and the uronic acid is often epimerised to iduronate, which may be 2-O sulfated (Xu and Esko, 2014, Ori et al., 2008, Murphy et al., 2004).

At least 883 extracellular regulatory proteins bind to transition and S-domains (Ori et al., 2008, Xu and Esko, 2014, Nunes, 2015). At the molecular level, analysis of the structural basis of the interaction of individual proteins with heparan sulfate and model polysaccharides (derivatives of the related heparin) show that there is a clear selectivity by proteins for particular patterns of sulfation (Maccarana et al., 1993, Faham et al., 1996, Jemth et al., 2002). At the tissue level, clear differences in the expression of sulfated sugar structures have been demonstrated, which impact on cell communication in development, homeostasis and disease (Theodoraki et al., 2015, Ford-Perriss et al., 2002, Thompson et al., 2011, Rudland et al., 1993, Chang et al., 2000, Friedl et al., 1997).

One important functional consequence of proteins binding heparan sulfate is its potential to control the movement of effectors between cells. Originally, endothelial cell extracellular matrix was demonstrated to be capable of storing fibroblast growth factor (FGF) 2, which could then transfer to its cellular receptors to stimulate the cells (Vlodavsky et al., 1987). Later, heparan sulfate in extracellular matrix was shown to control the diffusion of FGF2 (Dowd et al., 1999), which indicated that heparan sulfate had the potential to shape FGF2 gradients (FGF2 being both a growth factor and a morphogen (Serls et al., 2005)). Subsequently, the binding of a number of morphogens to heparan sulfate was shown to control their diffusion in contexts ranging from *Drosophila* to vertebrates (Makarenkova et al., 2009, Harada et al., 2009, Qu et al., 2012, Belenkaya et al., 2004, Izvolsky et al., 2003, Kleinschmit et al., 2010). However, this may not be universal (Zhou et al., 2012, Kicheva et al., 2007, Panakova et al., 2005). Moreover, it is not clear whether it is the selectivity of an effector for particular structures in the polysaccharide or just non-selective ion-exchange protein-polysaccharide interactions (Kreuger et al., 2006) that are important in regulating the effector's diffusion. A related issue is that heparan sulfate in extracellular matrix has been viewed as homogenous, that is there is no variation in the distribution of binding sites below the scale of tissue compartments. However, work with nanoparticle-labelled FGF2 demonstrated that the distribution of its binding sites in fibroblast pericellular matrix is heterogeneous and clustered from length scales of ~20 nm to 1 µm and above (Duchesne et al., 2012). Recently, biophysical experiments have shown that some effectors that bind heparan sulfate can cross link the chains of heparan sulfate (Migliorini et al., 2015). This suggests that heparan sulfate chains in extracellular matrix may be organised into supramolecular structures, which could impose the selectivity on protein binding that is of higher spatial order than possible with individual chains.

To test these ideas, we have used five FGFs, FGF1, 2, 6, 10 and 20, with well-characterised heparan sulfate binding sites and binding selectivity for structures in the polysaccharide (Xu et al., 2012, Li et al., 2015). These FGFs were expressed as N-terminal HaloTag fusions (Halo-FGFs) (Sun et al., 2015), which permitted specific fluorescent labelling. Measurement of the binding and diffusion of the Halo-FGFs in the pericellular matrix of fibroblasts revealed that there were very substantial differences between these FGFs in their level of binding, their spatial

distribution, and their diffusion. These data indicate that heparan sulfate chains in pericellular matrix are organised over length scales far greater than that of a single chain and that this serves to present distinct numbers and spatial patterns of binding sites for effectors, which in turn modulates the diffusion of the proteins.

Materials and methods:

Protein production

The FGFs and Halo-FGFs were produced exactly as described in detail previously (Xu et al., 2012, Sun et al., 2015). HaloTag protein was produced by digestion of Halo-FGF20 with TEV protease and purified by anion-exchange on DEAE Sepharose Fast Flow (GE Healthcare, Buckinghamshire, UK). Protein concentrations were determined by measuring their absorbance at 280 nm using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Leicestershire, UK). The biological activity of purified FGFs was detected by analysis of phosphorylation of $p44/42^{MAPK}$, as described in Section 3.2.

Protein labelling

HaloTag and Halo-FGFs (0.5 µM) were incubated with 2.5 µM HaloTag TMR ligand (Promega UK Ltd, Hampshire, UK) in 100 µL phosphate-buffered saline (PBS: 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 and 0.15 M NaCl, pH 7.4) at room temperature for 30 minutes, then kept on ice before use the same day. To determine the extent of labelling, TMR dye labelled Halo-FGFs were loaded onto a mini heparin agarose (BioRad, Hertfordshire, UK) column (20 µL) and washed with PBS containing 0.05% (v/v) Tween-20. The bound TMR labelled Halo-FGFs were eluted with 2 M NaCl buffered with phosphate (PB: 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). The quantum yields were measured in a fluorescence spectrophotometer (Varian, Walton-on-Thames, UK) by excitation at 561 nm and emission from 565 nm to 700 nm.

Cell culture

Rat mammary (Rama) 27 fibroblasts were cultured with Dulbecco's modified Eagle's medium (DMEM, Life Technologies, Paisley, UK) supplemented with 10%

(v/v) fetal calf serum (FCS, Labtech International Ltd, East Sussex, UK), 4 mM L-glutamine (Life Technologies), 0.75 % sodium bicarbonate (Life Technologies), 50 ng/mL insulin (Sigma-Aldrich, Dorset, UK) and 50 ng/mL hydrocortisone (Sigma-Aldrich), as described (Rudland et al., 1984).

Cell labelling

Rama 27 cells were cultured on glass bottomed imaging dished (CELLview™ Culture dish: 35mm non-treated glass bottom, Greiner Bio-one, Stonehouse, UK) and fixed with 4% (w/v) paraformaldehyde (PFA) dissolved in PBS. The fixed cells were washed with PBS three times and then incubated with 2 mL PBS containing 10 mg/ml BSA to block any remaining partially active fixative. The blocking medium was discarded after 15 minutes and the fixed cells were incubated with 1.5 mL 10 nM TMR dye, 2 nM TMR labelled HaloTag or 2 nM TMR labelled Halo-FGFs for 30 minutes at 37°C. The excess TMR dye and TMR dye labelled Halo-FGFs (TMR-Halo-FGFs) were removed by three washes with PBS. In competition experiments, the competitor was added along with the labelled Halo-FGF at concentrations indicated in the figure legends. Degradation of heparan sulfate was achieved in fixed Rama 27 cells by incubation with 1 mL heparinase I, II and III (50 mU/mL each in 100 mM sodium acetate and 0.1 mM calcium acetate, pH 7.0; gift from Prof. Jerry Turnbull, University of Liverpool). Chondroitin sulfate (including dermatan sulfate) was degraded by incubation with 1 mL chondroitinase ABC (Sigma-Aldrich; 400 mU/mL in PBS). In both cases, cells were incubated with the enzymes overnight at 37°C prior to incubation with Halo-FGFs.

Microscopy and imaging

A LSM780 confocal microscope with an environmental control chamber (Zeiss, Jena, Germany) was used to acquire cell imaging data with a 561 nm excitation filter and a 635 nm emission filter for detection. For all cell imaging, a 63X oil immersion lens (Plan-Apochromat 63x 1.4 oil DIC M27) was used. Cell images (67.3 µm x 67.3 µm, 512 x 512 pixels, 16 bits) containing bright field and the red fluorescence channel were collected for the binding assays. The flat cells (not dividing cells) were selected for imaging. Images were collected using identical microscope settings. To test whether the Halo-FGFs in Rama 27 pericellular matrix might be released into the

bulk medium from pericellular matrix, 10 consecutive images at 30 s intervals were acquired for Halo-FGF1, Halo-FGF2 and Halo-FGF6.

Fluorescence recovery after photobleaching (FRAP)

The fixed cells labelled with TMR-Halo-FGF1 (2 nM and 1 nM), TMR-Halo-FGF2 (2 nM), TMR-Halo-FGF6 (2 nM) and TMR-Halo-FGF10 (2 nM) were used for the FRAP experiments. The measurements were performed at 37°C. A square area (22.49 µm x 22.49 µm, 256 x 256 pixels, 8 bits) by the edge of the cell, very flat and no affect by the thick cell nucleus, was imaged 6 times with the 63X oil immersion lens and then the selected 2.5 µm (radius) disk area was bleached with the 561 nm laser at full power for 8 iterations (0.64 s in total). After that, another 195.6 s of images were acquired to measure the fluorescence recovery. An area free of cells and a non-bleached area on the same cell were selected to determine the background (subtracted in quantifications) and correct the photobleaching caused by the excitation laser during imaging, respectively. The fluorescence intensities of these three selected areas from 0 s to 197.2 s were extracted using ZEN 2012 software (Zeiss) for further analysis.

Data analysis

Fluorescence intensity of the labelled cells: the cell edges were automatically identified by using published Matlab codes (MathWorks, 2014) and the fluorescence intensities were averaged for each cell. The cell edges of low fluorescence labelled cells were detected in the bright field channel image (supplementary figures) and high fluorescence labelled cells were detected in the fluorescence channel image. The Matlab program for cell edge detection can be downloaded from GitHub (https://github.com/hscsun/DrawCellEdges.git) and the details of the program and instruction are included in Section 5.4.

FRAP data analysis: The background fluorescence intensity (I_b) was subtracted from both the bleached area (I) and non-bleached reference area (I_r). The photobleaching was corrected by the reference area and I_c is the corrected fluorescence intensity of the bleaching area. The fluorescence curves were not analysed by fitting with the exponential models, because many curves did not fit to this simple model, which

may be related to the complicated diffusion (*e.g.*, slow and fast) of FGFs in the fixed Rama 27 fibroblast pericellular matrix. A suitable model is being developed by colleagues from Mathematics department (University of Liverpool) to explain and understand the diffusion of FGFs in this matrix. Moreover, the fitting model is useful to extract the half recovery time and final recovery level of noisy fluorescence recovery curves, so it is not necessary for the collected smooth fluorescence recovery curves.

Eq 1: $I_c = (I - I_b)*((I_r[1\text{-}6] - I_b) / (I_r - I_b))$

Note: $I_r[1\text{-}6]$ means the averaged fluorescence intensity of the reference area of the first six images, which are used to get a stable fluorescence intensity for before bleaching; the other fluorescence intensities (I, I_b, I_c, I_{cn} and I_{dcn}) are applied to any image in the frame, but they are corresponded to the same image number in both sides of the equation for each calculation (from frame 1 to 1000 in this FRAP experiment).

The fluorescence intensity of bleaching area was normalised to the fluorescence intensity of the first image, where I_{cn} is the corrected and normalised intensity of the image.

Eq 2: $I_{cn} = I_c / I_c[1\text{-}6]$

Note: $I_c[1\text{-}6]$ means the average of the first six corrected fluorescence intensities from Eq 1.

To compare fluorescence recovery curves, the corrected and normalised fluorescence intensity of the first bleached image, $I_{cn}[7]$, was subtracted from the corrected fluorescence intensity of bleached area and the FRAP curve was normalised again, as in Eq 2.

Eq 3: $I_{dcn} = (I_{cn} - I_{cn}[7]) / (1 - I_{cn}[7])$

Note: [7] means the seventh image (or the first image after bleaching) for the normalised and corrected fluorescence intensity.

The final recovery level (I_f), the fluorescence intensity for the last measurement, and half recovery time ($\tau_{1/2}$) were extracted from the corrected and normalised curve acquired from Eq 3 by:

Eq 4: $I_f = I_{dcn}(\tau = \infty)$

Eq 5: $\tau_{1/2} = (\tau_a + \tau_b) / 2$

Note: ∞ means positive infinity and is equal to the time at which the last image was acquired in the actual experiments. τ_a is the time corresponding to the maximum value of the fluorescence intensities smaller than half of I_f; τ_b is the time corresponding to the minimum value of the fluorescence intensities larger than half of I_f.

The radial profiles of the bleaching area were extracted using a published Matlab code (Chattrapiban, 2007) and the photobleaching was corrected for each analysed image as described in Eq 1. The Matlab program for FRAP data analysis can be downloaded from GitHub (https://github.com/hscsun/ImagingDataAnalyzerForFRAP.git) and the details of the program and instruction are included in Section 5.4.

All calculations and image montages were done with Matlab R2014a.

Boxplots of the half recovery time and final recovery level of different Halo-FGFs were prepared in OriginPro 9 and the statistical analysis was processed with *Tukey* test in OriginPro 9. The data plot with standard deviation area was prepared using a published Matlab code (Campbell, 2010).

Results and Discussion:

Labelling Halo-FGFs with TMR-Halo ligand dye

The N-terminal HaloTag fusion does not affect the binding of FGFs to heparin or their biological activity and they are efficiently expressed (Chapter 3 and (Sun et al., 2015)). Thus, they provide a convenient means to prepare genetically encoded fluorescently labelled FGFs, whose excitation and emission properties can be altered by changing the HaloTag ligand (Los et al., 2008). We first tested whether the

HaloTag TMR ligand dye interacted with heparin or grossly affected the interaction of the FGFs with heparin. A mixture of HaloTag and a 5-fold excess of Halo-TMR dye was incubated for 30 minutes and loaded onto a mini heparin column. After three 50 µl washes with PBS containing 0.05% (v/v) Tween-20 (PBST) to remove the unbound dye, there was no red fluorescence detectable on the heparin column (Fig. 1 A). This indicated that neither HaloTag nor the TMR-Halo ligand dye bound to heparin. In contrast, the heparin column loaded with TMR labelled Halo-FGF2 gave strong red fluorescence (Fig. 1 A), which demonstrated that Halo-FGF2 was labelled with TMR-Halo ligand dye and retained its heparin-binding properties. Following purification of fluorescent dye labelled Halo-FGFs on mini heparin columns, the bound TMR-Halo-FGFs were eluted with 2 M NaCl. The fluorescence emission curves of the purified Halo-FGFs demonstrated that the emission peak of TMR dye remained at 580 nm and that the emission curves of these Halo-FGFs were quite similar (Fig. 1 B). Although there was a small difference of the fluorescence intensity for each TMR-Halo ligand labelled Halo-FGF, the results indicate that the labelling efficiency of the different Halo-FGFs was relatively consistent.

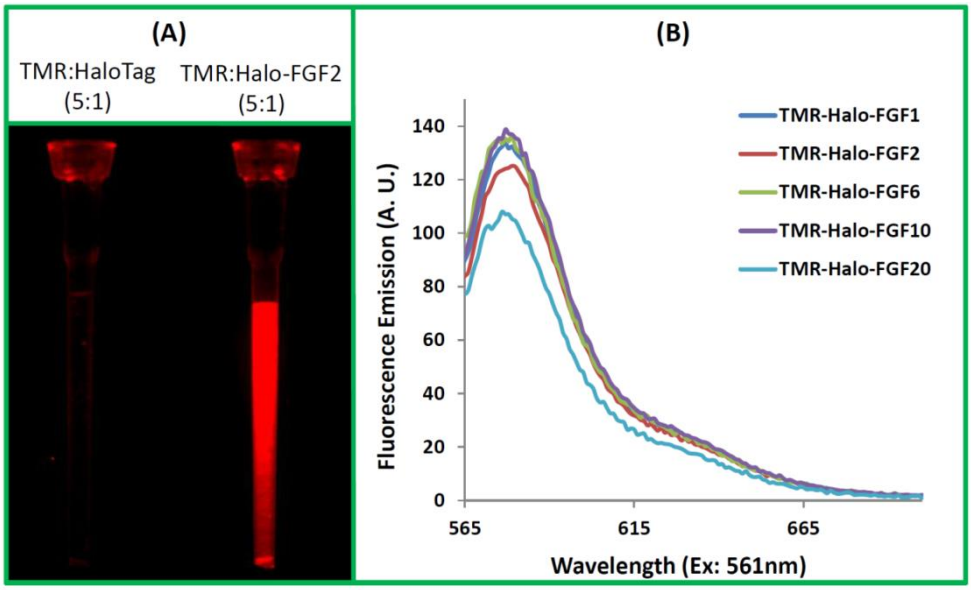


Figure 1: Conjugation and quantification of TMR dye labelled Halo-FGFs. Halo-TMR dye was used to label HaloTag and Halo-FGFs at a ratio of 5:1 (mole/mole). The labelled HaloTag and Halo-FGFs were loaded onto a mini heparin column, which was subsequently washed with PBST. **(A):** The HaloTag and Halo-FGF2 loaded heparin columns were visualised under a red fluorescence filter

(ImageQuant LAS 4000 imager, GE Healthcare). **(B):** The five TMR-labelled Halo-FGFs were loaded onto mini heparin-affinity chromatography columns, washed with PBS buffer and eluted with 2 M NaCl in the same buffer. The fluorescence intensities of the five purified Halo-FGFs were quantified in a fluorimeter by measuring the emission from 565 nm to 700 nm excited with 561 nm.

Binding of different Halo-FGFs to Rama 27 fibroblast pericellular matrix heparan sulfate

Rama 27 fibroblasts were fixed with paraformaldehyde prior to imaging, which stopped cellular biochemical processes. So, binding of FGFs to pericellular matrix will not be affected by internalisation. Since fixation is performed in an isosmotic buffer, such crosslinking will not alter the spatial relationship of the constituents, as might happen, for example, with a fixative that precipitates macromolecules, such as -20 ^{o}C methanol.

Halo-FGF2 The fixed Rama 27 fibroblasts were incubated with 2 nM Halo-FGFs to determine if their binding capacities to heparan sulfate in the pericellular matrix of these cells differed. Halo-FGF2 strongly bound to Rama 27 fibroblasts (Fig. 2 A). There was evidence for the heterogeneities in the distribution of the Halo-FGF2, evidenced by bright spots (Fig. 2 A arrows).

The BSA blocked cell dish was visualised by confocal microscopy. The cell edges were detected by the brightfield image (Fig. S1 B), and no autofluorescence from the cells was observed in the fluorescence channel image (Fig. 2 B). Using the same microscope settings, when TMR-Halo ligand alone and TMR-Halo ligand labelled HaloTag were incubated with the fixed Rama 27 cells, the fluorescence was the same as observed with a BSA blocked culture dish with cells; no red fluorescence was detectable (Figs 2 C-D). When the fixed cells were incubated with 2 nM TMR-labelled Halo-FGF2 and with either of two unlabelled competitors, 8 µM unlabelled FGF2 or Halo-FGF2, the binding was reduced to undetectable levels (Figs 2 E-F). These data indicate that non-specific binding of TMR-Halo ligand and of TMR-Halo ligand labelled HaloTag protein was within the levels of background fluorescence and that the fluorescence observed with labelled Halo-FGF2 in Rama 27 pericellular matrix (Fig. 2 A) was entirely due to the FGF2 moiety of the Halo-FGF2.

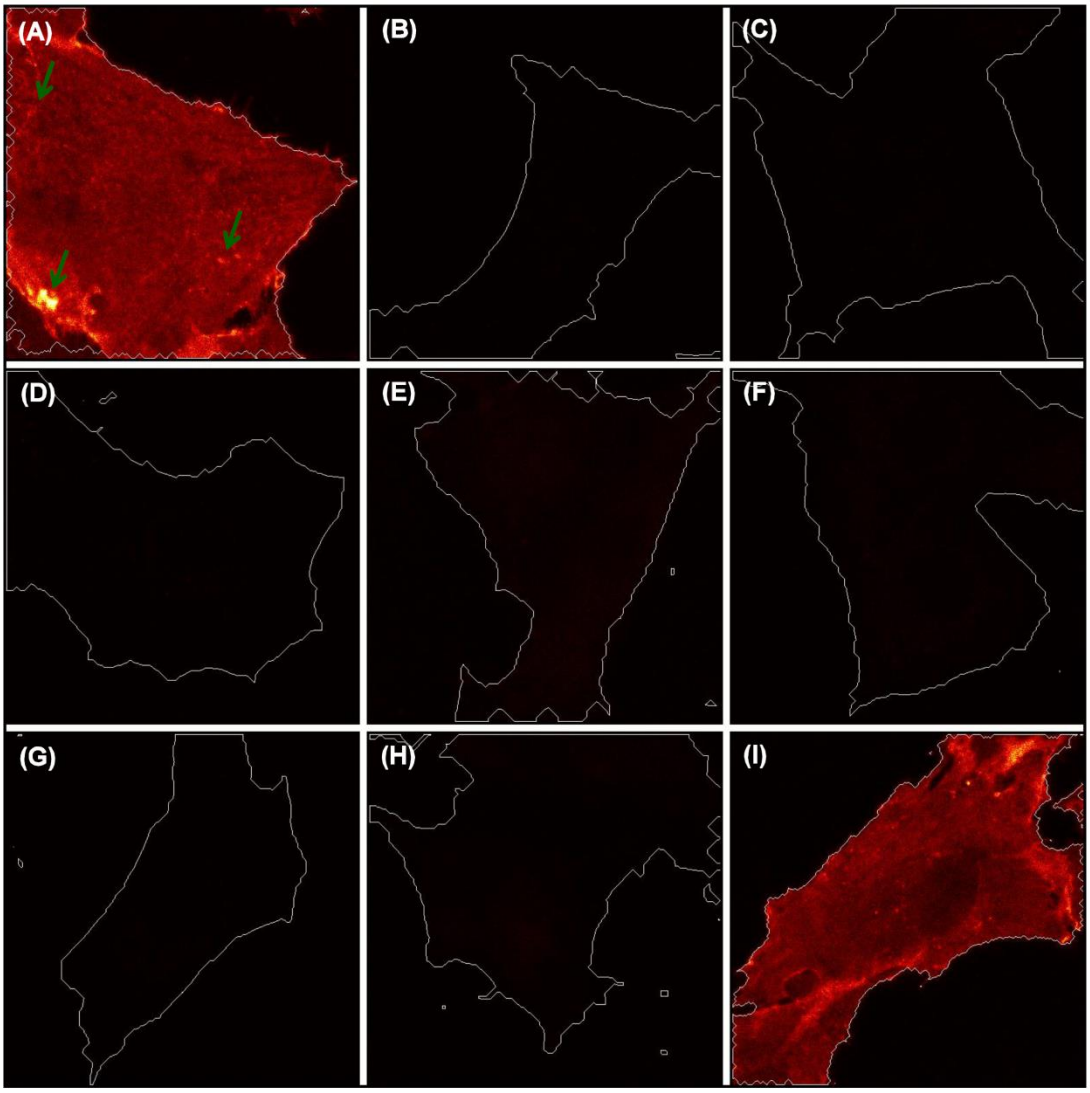


Figure 2: Halo-FGF2 binding to Rama 27 fibroblasts. The binding specificity of Halo-FGF2 to Rama 27 cells was determined by competing with unlabelled FGFs and heparin, and by digestion with heparinases and chondroitinase ABC. TMR-Halo-FGF2 (2 nM), Halo-TMR dye or TMR-HaloTag were used to label fixed Rama 27 fibroblasts for 30 min. The excess Halo-FGF2, Halo-TMR dye or TMR-HaloTag was removed by washing with PBS three times. The cell edges are highlighted with white lines. **(A):** TMR-Halo-FGF2 (2 nM); **(B):** unlabelled cells imaged to show the auto-fluorescence; **(C):** Halo-TMR dyes (2 nM), to measure the non-specific binding of ligand dye to cells or glass dish; **(D):** TMR-HaloTag (2 nM) to determine the level of binding of HaloTag. **(E and F):** Cells incubated with 2 nM TMR-Halo-FGF2 and 8 µM unlabelled Halo-FGF2 (E) or 8 µM FGF2 (F); **(G):** 2 nM TMR-Halo-FGF2 and 4 µg/mL heparin; **(H):** Cells were incubated with heparinases I, II and III to remove heparan sulfate and then incubated with 2 nM TMR-Halo-FGF2; **(I):** Cells

were incubated with chondroitinase ABC to digest chondroitin sulfate and then incubated with 2 nM TMR-Halo-FGF2. The corresponding bright field images are presented in Fig. S1. Image size: 67.3 µm X 67.3 µm.

To determine what Halo-FGF2 was binding to in the pericellular matrix of Rama 27 fibroblasts, a series of competition and enzyme digestion experiments were performed, again using the same microscope settings. Competition with heparin (4 µg/mL added with Halo-FGF2) abolished binding and fluorescence was reduced to background levels (Fig. 2 G). This indicates that FGF2 is likely bound to glycosaminoglycans of the pericellular matrix. Moreover, while heparin will effectively compete for binding of FGF2 to glycosaminoglycans, it still enables FGF2 to bind to the FGFR on these cells (Rahmoune et al., 1998). Therefore, Rama 27 fibroblasts were subjected to heparinase and chondroitinase ABC digestion to ascertain its binding partner(s). Incubation of fixed Rama 27 fibroblasts with heparinases I, II and III prior to the addition of Halo-FGF2 reduced the level of fluorescence to background levels (Fig. 2 H). In contrast, chondroitinase ABC digestion of the cells did not appreciably alter the binding of TMR-Halo-FGF2 to Rama 27 cell (Fig. 2 I). These data demonstrate that TMR-Halo-FGF2 is primarily bound to heparan sulfate in the pericellular matrix of Rama 27 fibroblasts. Moreover, these results are consistent with previous data, which indicate that more than 99% of binding sites for FGF2 on Rama 27 fibroblasts are heparan sulfate, and the FGFR less than 1% (Duchesne et al., 2012, Fernig et al., 1990).

Halo-FGF1 The binding of TMR labelled Halo-FGF1 to Rama 27 cells was somewhat stronger than that observed for Halo-FGF2 (Figs 2 A and 3 A). Since the labelling efficiencies of the Halo-FGFs are similar, this indicates that FGF1 at this concentration possesses more binding sites on these cells than FGF2. As for Halo-FGF2 the distribution of the fluorescence was not homogenous (Fig. 3 A). The lower fluorescence intensity in the centre of the cell was the result of the high focal plane of the plasma membrane in this region due to the underlying cell nucleus. The same competition and enzyme digestion experiments performed with Halo-FGF2 were done with Halo-FGF1, to identify its binding partner(s) in Rama 27 pericellular matrix. Both unlabelled 8 µM Halo-FGF1 and FGF1 effectively competed with 2 nM TMR-Halo-FGF1 (Figs 3 B and C). Addition of 4 µg/mL heparin with TMR-Halo-

FGF1 also abolished detectable binding of the latter to Rama 27 fibroblasts (Fig. 3 D). Treatment of fixed Rama 27 cells with heparinases was similarly effective in reducing the binding of TMR-Halo-FGF1 below the limit of detection (Fig. 3 E). However, digestion with chondroitinase ABC increased the level of fluorescence (Fig. 3 F). The increase in binding of Halo-FGF1 observed after chondroitinase ABC treatment may indicate that removal of chondroitin sulfate changed the structure of ECM and somehow increased the number of available heparan sulfate binding sites for FGF1 (Fig. 3 F). Collectively, these data demonstrate that the detectable fluorescent Halo-FGF1, like the Halo-FGF2, is bound to the heparan sulfate of the pericellular matrix of Rama 27 fibroblasts. Though FGF1 binds heparan sulfate preferentially, it also binds dermatan sulfate more weakly (Xu et al., 2012), but dermatan sulfate binding sites are either not available or too weak in Rama 27 pericellular matrix, since chondroitinase ABC treatment increased, rather than decreased binding. Interactions with the FGFR are below the level of detection, which is consistent with the relative numbers of binding sites corresponding to heparan sulfate and the FGFR established previously for FGF2 in these cells (Fernig et al., 1990, Duchesne et al., 2012).

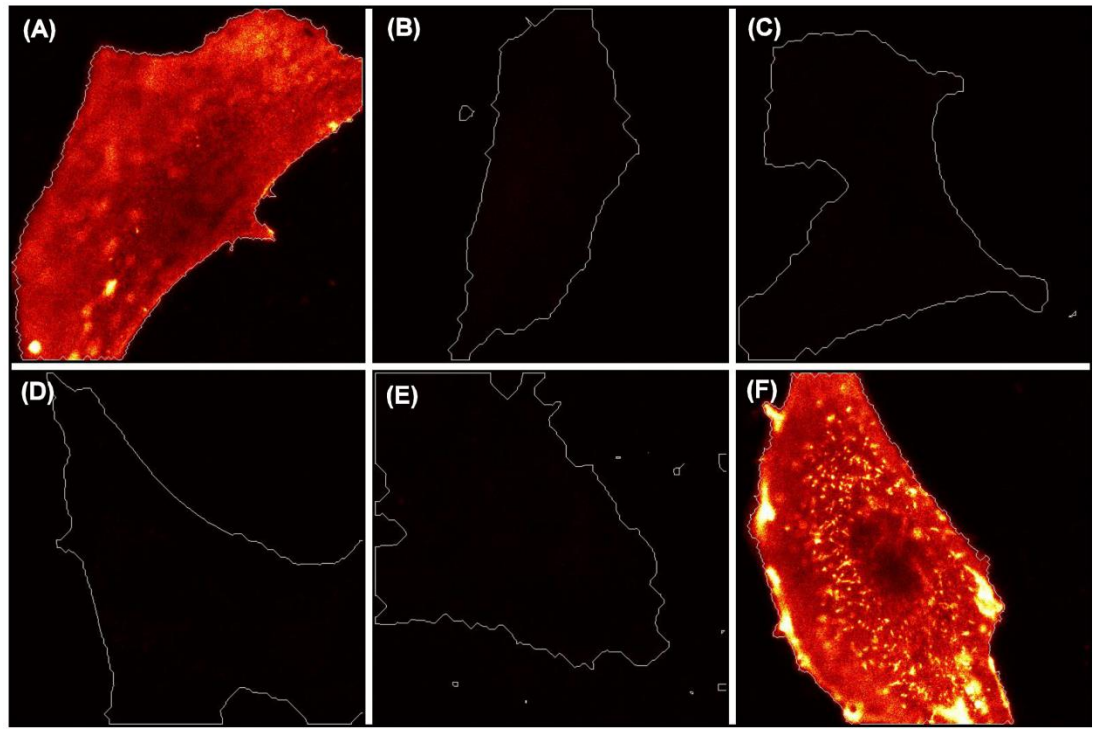


Figure 3: Halo-FGF1 binding to Rama 27 fibroblasts. The binding specificity of Halo-FGF1 to fixed Rama 27 fibroblasts was tested by competing with unlabelled FGF1 and heparin and by digestion with heparinase and chondroitinase ABC. **(A):** Cells were incubated with 2 nM TMR-Halo-FGF1 at 37^{o}C for 30 min; **(B and C):** 2 nM TMR-Halo-FGF1 was added with 8 µM unlabelled Halo-FGF1 (B) or 8 µM unlabelled FGF1 (C); **(D):** TMR-Halo-FGF1 (2 nM) in the presence of 4 µg/mL heparin; **(E and F):** TMR-Halo-FGF1 binding to Rama 27 fibroblasts previously subjected to digestion with heparinase I, II and III, and chondroitinase ABC, respectively. The corresponding bright field images are presented in Fig. S2. Image size: 67.3 µm X 67.3 µm.

Halo-FGF6 Halo-FGF6 bound only slightly less than FGF2 to fixed Rama 27 fibroblasts and again the fluorescence was not homogenous (Fig. 4 A). No binding of Halo-FGF6 was observed on the fixed Rama 27 fibroblasts when TMR-Halo-FGF6 was added with 4 µg/mL heparin (Fig. 4 B). Similar to FGF1 and FGF2, digestion of heparan sulfate by heparinase decreased the binding of TMR-Halo-FGF6 to undetectable levels (Fig. 4 C), whereas digestion of chondroitin sulfate and dermatan sulfate led to an increase in Halo-FGF6 binding to the cells, as seen with FGF1 (Fig. 4 D). These results indicated that the detectable Halo-FGF6 was bound to heparan sulfate in the pericellular matrix of Rama 27 fibroblasts. The number of these sites is similar to those recognised by FGF2, but chondroitin sulfate (or dermatan sulfate) would appear to prevent directly or indirectly some Halo-FGF6 binding to the heparan sulfate in the pericellular matrix.

Halo-FGF10 Halo-FGF10 clearly only bound to some areas of the pericellular matrix, whereas in other areas, virtually no binding was detected (Fig. 4 E). Thus, the binding of Halo-FGF10 to Rama 27 fibroblasts was characterised by very substantial heterogeneities. The binding sites on Rama 27 fibroblasts for FGF10 were also blocked by addition of 4 µg/mL heparin, which effectively prevented FGF10 binding to fixed Rama 27 fibroblasts (Fig. 4 F). Digestion of heparan sulfate with heparinase I, II and III reduced the level of binding of Halo-FGF10 (Fig. 4 G), but unlike Halo-FGF1, Halo-FGF2 and Halo-FGF6, did not abolish it (Figs 2 H, 3 E and 4 C). Moreover, digestion of chondroitin sulfate / dermatan sulfate with chondroitinase ABC also reduced the amount of bound Halo-FGF10 (Fig. 4 H).

These results indicated that Halo-FGF10 may bind to both heparan sulfate and chondroitin sulfate in Rama 27 fibroblasts pericellular matrix. Therefore, a double digestion (heparinase and chondroitinase) was performed. When both sets of glycosaminoglycans were digested, the level of bound Halo-FGF10 was nearly undetectable (Fig. 4 I), demonstrating that FGF10 does indeed bind to both chondroitin (dermatan) and heparan sulfate.

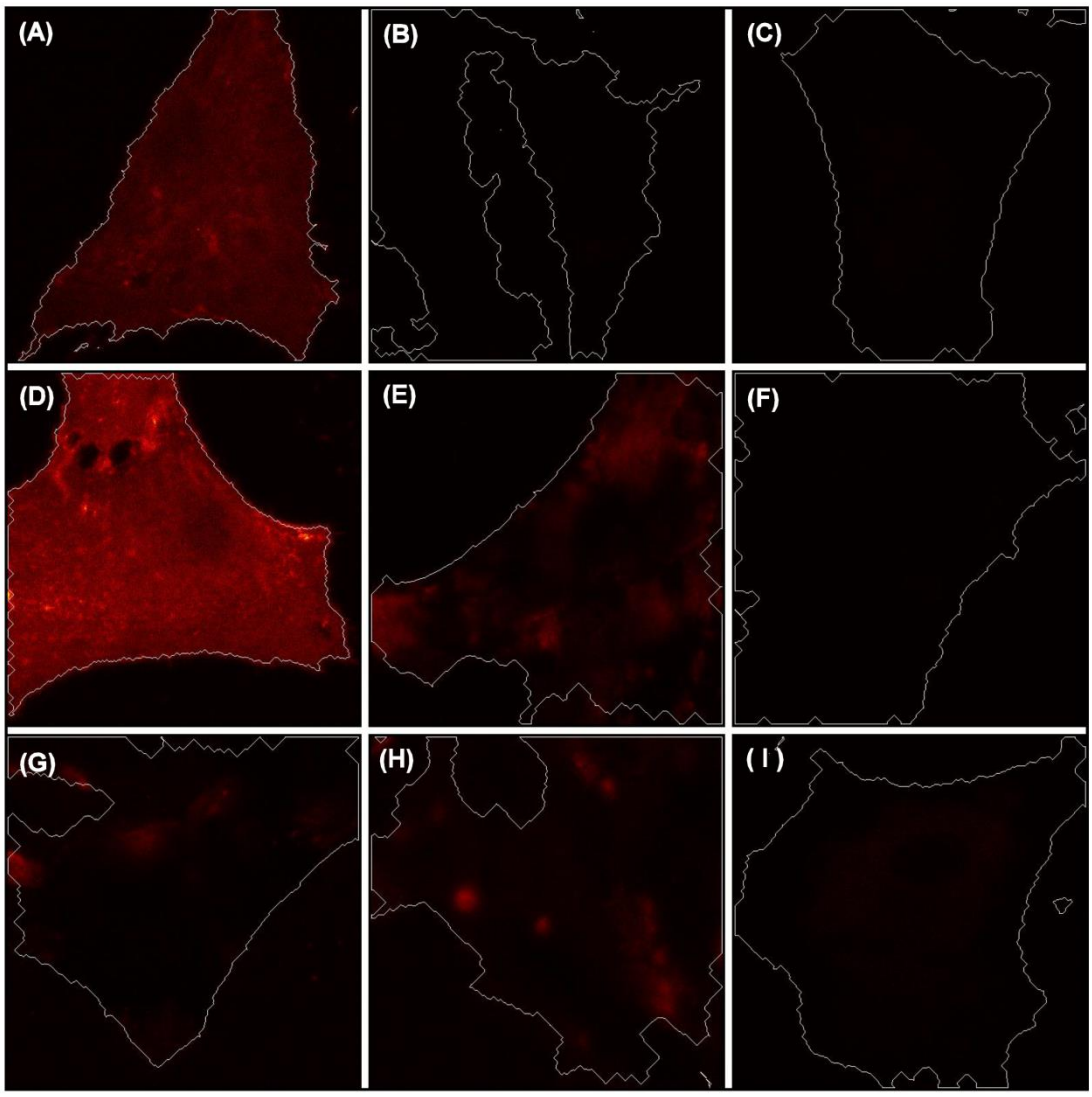


Figure 4: Binding specificity of Halo-FGF6 and Halo-FGF10 to Rama 27 fibroblasts. Halo-FGF6 and Halo-FGF10 were used to label fixed Rama 27 fibroblasts and heparin and the enzyme digested fibroblast pericellular matrix was used to determine their binding specificity. **(A):** Rama 27 fibroblasts were incubated with 2 nM TMR-Halo-FGF6; **(B):** 2 nM TMR-Halo-FGF6 in the presence of 4 µg/mL heparin; **(C and D):** 2 nM TMR-Halo-FGF6 binding to Rama 27 fibroblast

pericellular matrix digested with heparinase I, II and III, and chondroitinase ABC, respectively; **(E):** Rama 27 fibroblasts were incubated with 2 nM TMR-Halo-FGF10; **(F):** TMR-Halo-FGF10 (2 nM) in the presence of 4 µg/mL heparin; **(G and H):** TMR-Halo-FGF10 (2 nM) binding to Rama 27 fibroblasts pericellular matrix digested with heparinase I, II and III, and chondroitinase ABC, respectively; **(I):** 2 nM Halo-FGF10 binding to the pericellular matrix digested by both heparinase I, II and III, and chondroitinase ABC. The corresponding bright field images are presented in Fig. S3. Image size: 67.3 µm X 67.3 µm.

Comparison of binding of Halo-FGFs to Rama 27 cell pericellular matrix heparan sulfate

Quantification of the level of binding of the Halo-FGFs to Rama 27 cell pericellular matrix revealed some marked differences. The fluorescence intensity of the Halo-FGFs was determined by calculating the averaged fluorescence intensity of the highlighted cell area to compare their binding capacities to the pericellular matrix. There were more binding sites for Halo-FGF1 than the other Halo-FGFs (Fig. 5 A). Based on Tukey t-test, the binding capacities of Rama 27 pericellular matrix for Halo-FGF2 and for Halo-FGF6 were also significantly different ($p=0.005$, Tukey test) with Halo-FGF2 possessing more binding sites. In terms of binding intensity, Halo-FGF6 and Halo-FGF10 did not have significant difference ($p=0.08$), but the distributions of Halo-FGF10 and the other three Halo-FGFs were clearly not the same (Fig. 5 A). For example, Halo-FGF6 was more evenly distributed in pericellular matrix, with a level of inhomogeneity similar to that seen with Halo-FGF2, whereas Halo-FGF10 only bound to specific areas of the pericellular matrix. In contrast, Halo-FGF20 bound extremely weakly, if at all, which was not detectable (Figs S4 A and B).

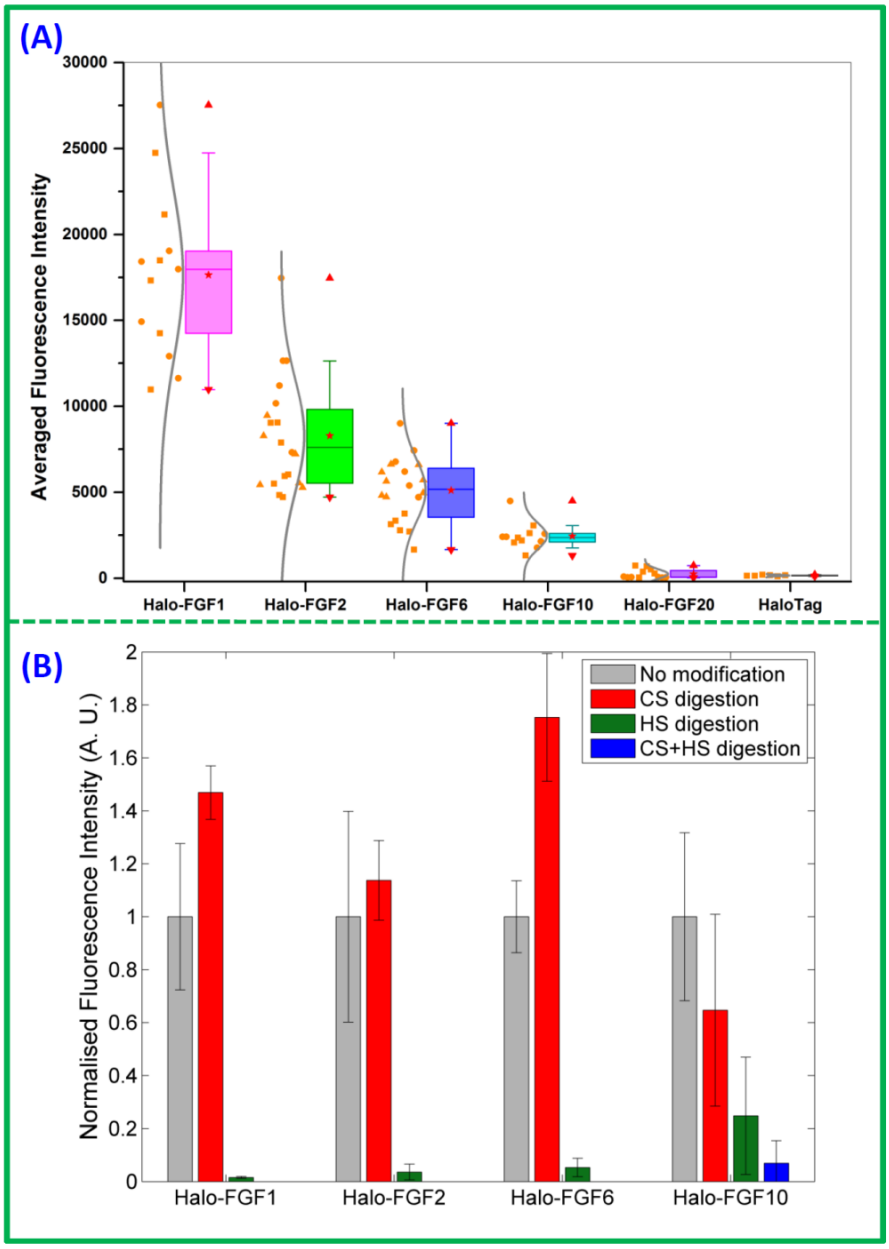


Figure 5: Quantification of binding of different Halo-FGFs to Rama 27 fibroblast pericellular matrix. (A): TMR labelled Halo-FGF1, Halo-FGF2, Halo-FGF6, Halo-FGF10, Halo-FGF20 and HaloTag (all 2 nM) were incubated with fixed Rama 27 fibroblasts, as described in Figs 2 A, 3 A, 4 A, 4 E, S4 A and S4 B. The fluorescence in the highlighted cell area was averaged to quantify the level of binding of the FGF to Rama 27 pericellular matrix. Fluorescence intensities on different cells in the same set of dish and different sets of dishes were acquired and are shown as a box plot. Each symbol corresponds to independent dishes of cells measured on different days; **(B):** The binding intensities of Halo-FGF1, Halo-FGF2, Halo-FGF6 and Halo-FGF10 to Rama 27 fibroblasts pericellular matrix digested

with heparinase I, II and III, and with chondroitinase ABC were quantified and normalised to the values obtained with untreated matrix.

These results differ from those obtained upon affinity chromatography of these FGFs to heparin (Sun et al., 2015). The previous work indicated that Halo-FGF2, Halo-FGF1 and Halo-FGF10 could bind to heparin in 0.6 M NaCl, and 1 M or higher NaCl was required to efficiently elute them from heparin-affinity chromatography matrices. Both Halo-FGF6 and Halo-FGF20 could stably bind to heparin in 0.4 M NaCl, though less Halo-FGF20 was bound than Halo-FGF6. These data highlight that binding to heparin, which is far more sulfated than heparan sulfate, does not reflect the binding capacity of heparan sulfate, which is both less sulfated and more structurally diverse (Ori et al., 2008, Xu and Esko, 2014), which allows a far more selective interaction with individual proteins (Jemth et al., 2002, Xu, 2012).

The binding and competition data demonstrate that the detectable binding of the four Halo-FGFs is to glycosaminoglycans in the pericellular matrix. In the presence of heparin these FGFs will interact with their receptor tyrosine kinase (Zhang et al., 2006b). Thus, the absence of binding of Halo-FGF2 detected with heparin is in agreement with previous work, which showed that the number of heparan sulfate binding sites for FGF2 is several orders of magnitude greater than the number of receptors (Duchesne et al., 2012). A similar difference is, therefore, likely to exist for FGF1, FGF6 and FGF10, since binding was not detected in the presence of heparin (Figs 2 H, 3 E, 4 C and 4 I). Whereas Halo-FGF1, Halo-FGF2 and Halo-FGF6 interacted only with heparan sulfate, Halo-FGF10 had a significant interaction with chondroitin sulfate (and/or dermatan sulfate) species on Rama 27 fibroblasts (Fig. 5 B). FGF1 has previously been shown to interact with dermatan sulfate, but not chondroitin sulfate, whereas FGF7, which is in the same subfamily as FGF10, interacts weakly with both chondroitin sulfate and dermatan sulfate (Xu et al., 2012). Consistent with the latter result, FGF10 has been shown to bind to chondroitin sulfate and dermatan sulfate (Li et al., 2015). In the case of Halo-FGF1, either the interaction with dermatan sulfate is too weak to be detectable or there is little dermatan sulfate with appropriate binding structures in Rama 27 cell pericellular matrix. In contrast, the interaction of Halo-FGF10 with chondroitin sulfate and/or dermatan sulfate on these cells is sufficiently strong to be detected (Fig. 5 B).

The increase in binding observed with Halo-FGF1 and Halo-FGF6 upon chondroitinase ABC treatment of cells suggests that chondroitin sulfate may somehow mask heparan sulfate binding sites for these Halo-FGFs. Whether such masking occurs directly or due to bridging by endogenous proteins that bind both chondroitin sulfate and heparan sulfate is not known. It is intriguing that the effect is not seen with Halo-FGF2, since this is in the same subfamily as FGF1 and the major difference in binding selectivity between these FGFs is that FGF1 readily binds tracts of sulfated saccharides containing 6-O sulfated glucosamine with one of N-sulfated glucosamine or 2-O sulfated iduronic acid, whereas FGF2 binds these very poorly (Uniewicz et al., 2010, Jemth et al., 2002, Maccarana et al., 1993). Thus, the masking effect of chondroitin sulfate on Halo-FGF1 binding to heparan sulfate may be related to its interactions with such structures in heparan sulfate.

The binding of all the Halo-FGFs was observed to be heterogeneous. The differences in binding between FGF10 and FGF1, FGF2 and FGF6 together with the complete absence of binding of FGF20 demonstrate that the observed heterogeneity of binding is not simply due to differences in the distribution of HS. Instead, these data indicate that it is the distribution of binding sites of polysaccharide chains for Halo-FGF1, Halo-FGF2 and Halo-FGF6 in heparan sulfate and for Halo-FGF10 in heparan sulfate and chondroitin sulfate that are not evenly distributed across the pericellular matrix. This is consistent with similar observation made with gold nanoparticle labelled FGF2 by photothermal microscopy (optical resolution) and by transmission electron microscopy (molecular resolution) (Duchesne et al., 2012, Nieves et al., 2015). Taken together, these data suggest that the previously observed clustering of FGF2 binding sites in heparan sulfate of Rama 27 cell pericellular matrix may be a more general phenomenon, since it is seen here with four FGFs from three different subfamilies that possess different binding selectivity for heparan sulfate (Li et al., 2015, Xu et al., 2012). This suggests that the binding sites for these FGFs are spatially organised in Rama 27 pericellular matrix and this is likely to extend to supramolecular length scales (distance equivalent to several/many heparan sulfate chains). Such organisation would arise from the interaction of heparan sulfate and (for FGF10) chondroitin sulfate / dermatan sulfate chains with their endogenous binding proteins, which for heparan sulfate have been catalogued to at least 883 (Ori et al., 2011, Nunes, 2015).

Detection of FGF diffusion by FRAP

The differences in the binding of Halo-FGF1, 2, 6, 10 and 20 to Rama 27 pericellular matrix, relate, at least in part, to differences in the structures these FGFs bind in heparan sulfate (and chondroitin sulfate/dermatan sulfate for FGF10). It is established in some cases that the interaction of proteins with heparan sulfate can control their movement in the extracellular space (Duchesne et al., 2012, Akiyama and Gibson, 2015, Muller et al., 2013, Dowd et al., 1999, Yan and Lin, 2009). Therefore, to determine if the differences in heparan sulfate binding may result in differences in movement in extracellular matrix, we measured the diffusion of Halo-FGF1, 2, 6 and 10 in Rama 27 pericellular matrix by FRAP.

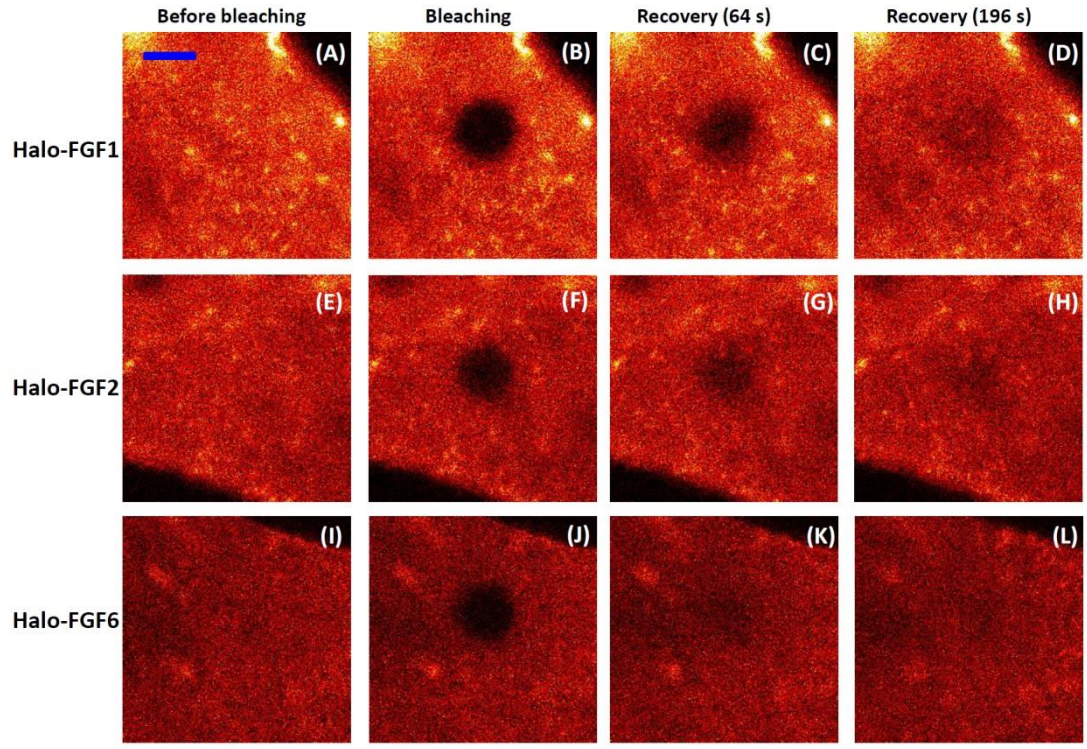


Figure 6: FRAP of Halo-FGF1, Halo-FGF2 and Halo-FGF6 in Rama 27 fibroblast pericellular matrix. Fixed Rama 27 fibroblasts were used to provide a pericellular matrix that could be probed with Halo-FGF1, Halo-FGF2 and Halo-FGF6. A 5 µm radius disk area on the cell was bleached by full power laser to measure the recovery of the fluorescence in the bleached area. **(A, E and I):** TMR-Halo-FGF1, -FGF2 and -FGF6 labelled cells before bleaching; **(B, F and I):** Same areas as (A, E and I), but following the bleaching of a 5 µm radius disk; **(C, G and**

K): The partial recovery of fluorescence in the bleached area 64 s after bleaching; **(D, H and L):** Images acquired when the bleached area had recovered to a stable level (196 s). Size of the scale bar: 5 µm.

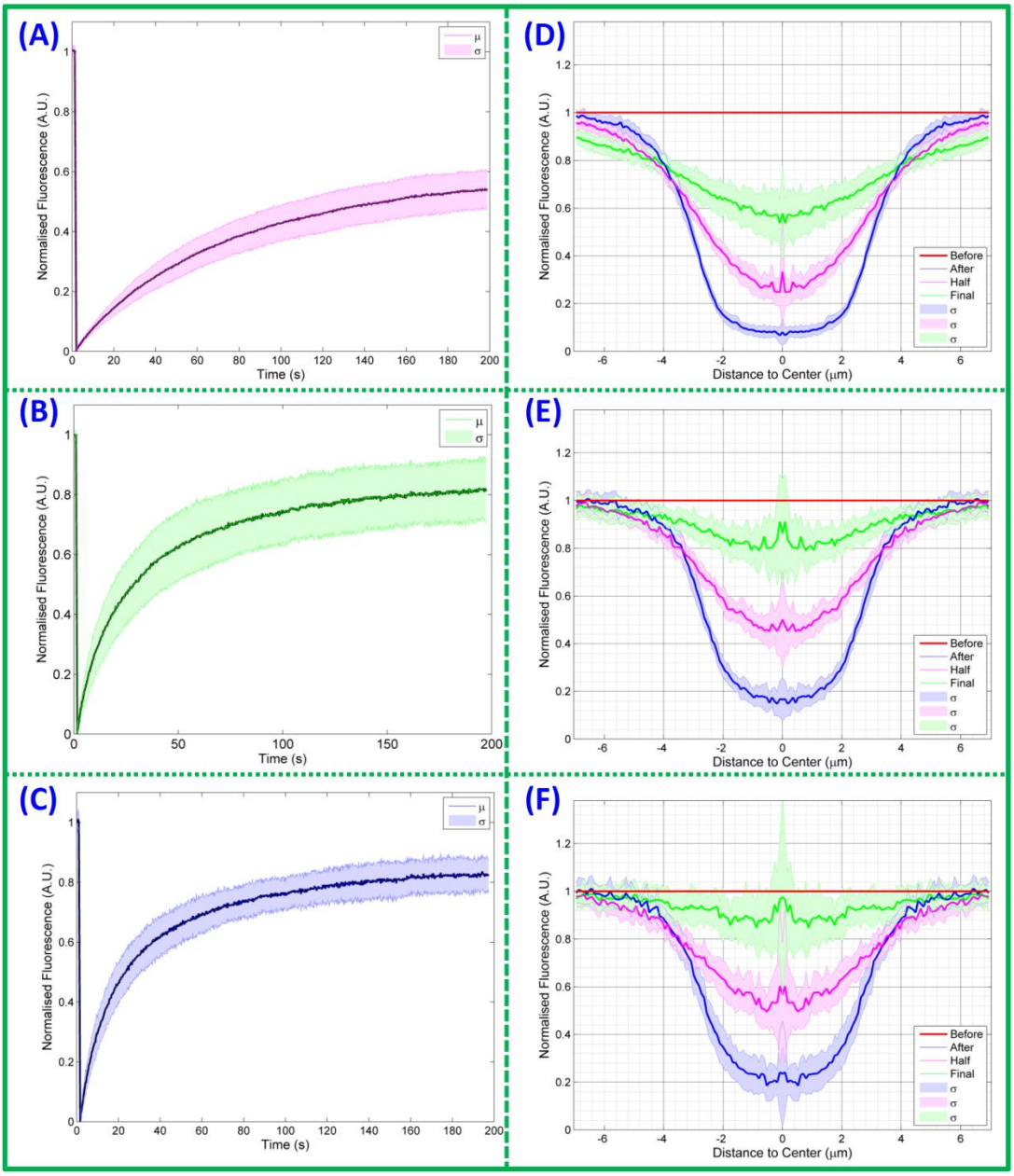


Figure 7: Fluorescence recovery curves and recovery radial profiles of Halo-FGF1, Halo-FGF2 and Halo-FGF6 in Rama 27 fibroblast pericellular matrix. The fluorescence intensity of the bleaching area was analysed, as described in materials and methods, to identify the different recovery patterns. The radial profile of the bleached area was extracted from the imaging data to reflect how the FGFs exchanged between the bleached area and the surrounding non-bleached pericellular

matrix. **(A, B and C):** The normalised fluorescence intensities of Halo-FGF1 (A), Halo-FGF2 (B) and Halo-FGF6 (C) in the bleached area were plotted against time (Average of 10 measurements for Halo-FGF1, 17 measurements for Halo-FGF2 and 28 measurements for Halo-FGF6); **(D, E and F):** The radial profiles of the bleached area before bleaching, immediately after bleaching, when fluorescence had reached half the final recovery value and at final recovery were extracted from the imaging data. Multiple repeats were applied to acquire the standard deviation. The mean of radial profiles for each FGF was plotted with standard deviation area against the distance to the centre of the bleached disk area (18 measurements for Halo-FGF1, 23 for Halo-FGF2 and 17 for Halo-FGF6). μ is the mean value of multiple fluorescence intensity curves for each FGF; σ is the standard difference; 'Before' is before bleaching; 'After' is the image immediately after bleaching; 'Half' is the time when the fluorescence was recovered to half of the final recovery level; 'Final' is the time for the last measurement.

The FRAP experiments employed the same labelling protocol as the imaging ones. Fixed cells were again used, because this allowed the measurement of the diffusion of each Halo-FGF in pericellular matrix to be made without any confounding effects that might have arisen due to the movement of cells or of membrane. The amine reactive fixative (paraformaldehyde) will not affect the binding of the Halo-FGFs in heparan sulfate, since these proteins recognise oligosaccharide structures where the amino group of the glucosamine residues is sulfated (Ori et al., 2008, Xu and Esko, 2014). However, the fixative may cross link endogenous multivalent heparan sulfate binding proteins and the core proteins of heparan sulfate proteoglycans. This may then restrict movement of heparan sulfate chains and diffusion in the membrane of the heparan sulfate proteoglycan core proteins both of which will restrict the freedom of the heparan sulfate chains (Duchesne et al., 2012), though this effect may be less pronounced on glycosyl-phosphatidylinositol anchored glypicans than transmembrane core proteins such as syndecans (Tanaka et al., 2010).

Another important feature of these experiments is that following the binding of Halo-FGFs to heparan sulfate in the pericellular matrix, the cells were washed to remove unbound Halo-FGF. Trapping of FGF2 on heparan sulfate in the extracellular matrix has been well documented (Duchesne et al., 2012, Zehe et al., 2006, Dowd et al.,

1999, Mundhenke et al., 2002, Taverna et al., 2003, Shute et al., 2004, Vlodavsky et al., 1987) and, given a suitable density of heparan sulfate binding sites, is a general property of extracellular matrix (Dowd et al., 1999, Lin, 2004, Yan and Lin, 2009, Belenkaya et al., 2004). As for FGF2 (Duchesne et al., 2012), Halo-FGF1, Halo-FGF2 and Halo-FGF6 bound to the pericellular matrix did not dissociate appreciably into the bulk culture medium over 270 s (Fig. S5). Thus, since FRAP measurement were made in 197 s, dissociation into the bulk culture medium followed by re-association with heparan sulfate in the pericellular matrix cannot contribute to the recovery of fluorescence. Instead, the recovery of fluorescence will be due to diffusion of these Halo-FGFs within the pericellular matrix.

After the bleaching iterations, the selected area became dark (Figs 6 A, B, E, F, I and J). Recovery of fluorescence then occurred (Figs 6 C, D, G, H, K and L). These data demonstrate that the Halo-FGFs were able to diffuse between the bleached and surrounding areas of pericellular matrix in fixed Rama 27 cells. Thus, while these Halo-FGFs were clearly trapped within the pericellular matrix (Fig. S5), they were able to diffuse within it. Movement of nanoparticle labelled FGF2 has similarly been evidenced before by photothermal imaging, tracking and raster image correlation spectroscopy (Nieves et al., 2015, Duchesne et al., 2012). Earlier work using diffusion chamber also demonstrated that FGF2 trapped on heparan sulfate in extracellular matrix was mobile (Dowd et al., 1999). Thus, the present data demonstrate that the movement of proteins bound to heparan sulfate and trapped in extracellular matrix is likely to be a more general phenomenon.

Quantification of diffusion of Halo-FGF1, Halo-FGF2, Halo-FGF6 and Halo-FGF10

The fluorescence intensity of the bleached area during recovery was quantified as the normalised fluorescence (Materials and Methods). In the case of Halo-FGF1, recovery was partial after 64 s and still not complete by 196 s (Figs 6 C-D and videos S7 and S8). The fluorescence recovery curve shows that Halo-FGF1 fluorescence in the bleached area recovered relatively slowly and by 196 s only half the fluorescence was recovered (Fig. 7 A). The decrease of fluorescence intensity of the reference area was due to the photobleaching by the imaging laser (Figs S6 A-B), since Halo-FGF1, Halo-FGF2 and Halo-FGF6 could be trapped in the pericellular matrix for

more than 4.5 min (Fig. S5), as discussed above. The recovery of fluorescence was greater for Halo-FGF2, although the bleached area is still perceptible after 196 s (Figs 6 G- H and video S9). Quantification of the recovery of Halo-FGF2 fluorescence demonstrates that this is substantially faster than that of Halo-FGF1 and the final level of fluorescence, 80%, was higher (Fig. 7 B). The fluorescence of Halo-FGF6 recovered similar to that of Halo-FGF2 (Figs 6 K-L and video S10). The rate of fluorescence recovery of Halo-FGF6 was somewhat faster than Halo-FGF2, though the level of recovery attained after 196 s was similar (Figs 7 B-C). The weak photobleaching for Halo-FGF2 might suggest that the bleached Halo-FGF2 and Halo-FGF6 during imaging could be quickly exchanged into the surrounding areas (Figs S6 C-D).

The fluorescence recovery curves (Figs 7 A-C) allowed the calculation of the half recovery time, which is directly related to the movement of molecules in the FRAP experiments and the relative proportions of mobile and immobile Halo-FGF. The half recovery times demonstrated that Halo-FGF1 diffused more slowly in the pericellular matrix of Rama 27 fibroblasts than Halo-FGF2 or Halo-FGF6 (Fig. 8 A). Moreover, Halo-FGF6 had the shortest half recovery time (16 s), which was significantly ($p=0.0008$, Tukey test) faster than that of Halo-FGF2 (22 s) and clearly of Halo-FGF1 (49 s). Fluorescence recovery is due to the diffusion of the Halo-FGFs. Thus, the difference of final level of recovered fluorescence and the initial fluorescence is indicative of the fraction of immobile Halo-FGFs. Only 52% of Halo-FGF1 was mobile, whereas 81% of Halo-FGF2 and 82% of Halo-FGF6 were mobile (Fig. 8 B). Previous work demonstrates that FGF2 that appears immobile at the resolution of a confocal microscope will in fact be undergoing confined motion, diameter ~100 nm (Duchesne et al., 2012). Like Halo-FGF2, Halo-FGF1 and Halo-FGF6 are also bound to heparan sulfate in the pericellular matrix. Though there are clear differences in the distribution and number of their available binding sites on heparan sulfate chains, it seems reasonable to suggest that the immobile fraction of Halo-FGF1 and Halo-FGF-6 are also undergoing similar confined motion.

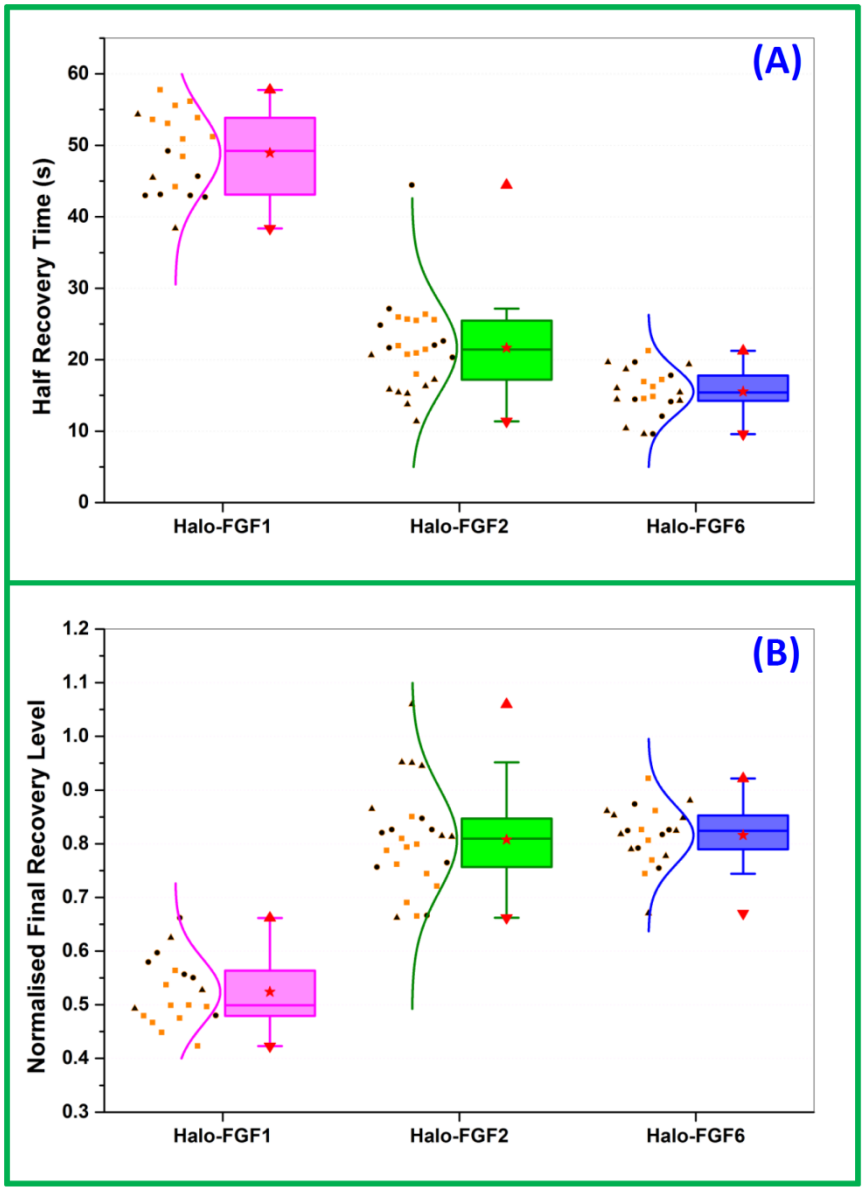


Figure 8: Quantification of moving speed and mobile/immobile fractions of Halo-FGF1, Halo-FGF2 and Halo-FGF6 on Rama 27 fibroblasts. The half recovery time and final recovery level were extracted from each fluorescence recovery curve, as described in materials and methods. **(A):** The half recovery times for Halo-FGF1, Halo-FGF2 and Halo-FGF6 were plotted to compare their diffusion speeds in the pericellular matrix. Each half recovery was extracted from one FRAP experiment, which shows the time it took to recover to half of the final fluorescence intensity in each fluorescence recovery curve; **(B):** Normalised final recovery levels of the three FGFs were used to determine the ratio of mobile and immobile FGF in the pericellular matrix. The final recovery level shows the mobile fraction and the immobile fraction is its difference from 100%.

Analysis of the movement of FGF2 at the single molecule level revealed that it undergoes different types of diffusive motion over different length scales. To see if some insight could be gained from the present average measurements of Halo-FGF diffusion into the types of movements the FGFs underwent, the fluorescence of the bleached area and surrounding unbleached area were determined as a series of radial profiles, diameter 14 µm. These analyses are presented as the radial profile at selected times: before bleaching, after bleaching, at the time corresponding to half recovery of the final fluorescence, and at final recovery. The results show that the radial profile after bleaching (Figs 7 D-F, blue lines) is "U" shaped, but, as the bleached area recovered, the profile (Figs 7 D-F, pink lines) it became more "V" shaped. Moreover, for Halo-FGF1, as the recovery profile of the bleached area (2.5 µm radius) increased, there was a small decrease in fluorescence in the surrounding unbleached area (Fig. 7 D, pink line and green line). Together, this suggests that the majority of the movement of the Halo-FGFs at these time scales is over 1 µm or less, corresponding to the confined and simple diffusive motion observed previously with FGF2 and that FGF1 may undergo comparatively little fast and directed diffusion (Duchesne et al., 2012). In contrast, the half recovery profiles of Halo-FGF2 and Halo-FGF6 (Figs 7 E-F, pink lines) were more "U" shaped and the fluorescence of the surrounding unbleached areas was not much affected during recovery (Fig. 7 D, pink line). Moreover, the final recovery profiles of Halo-FGF2 and Halo-FGF6 were close to that seen before bleaching (Figs 7 E-F, green lines). These data are consistent with the previous demonstration that FGF2 can undergo fast and directed diffusion in addition to confined and simple diffusive motion and it would appear that Halo-FGF6 may undergo similar types of movement.

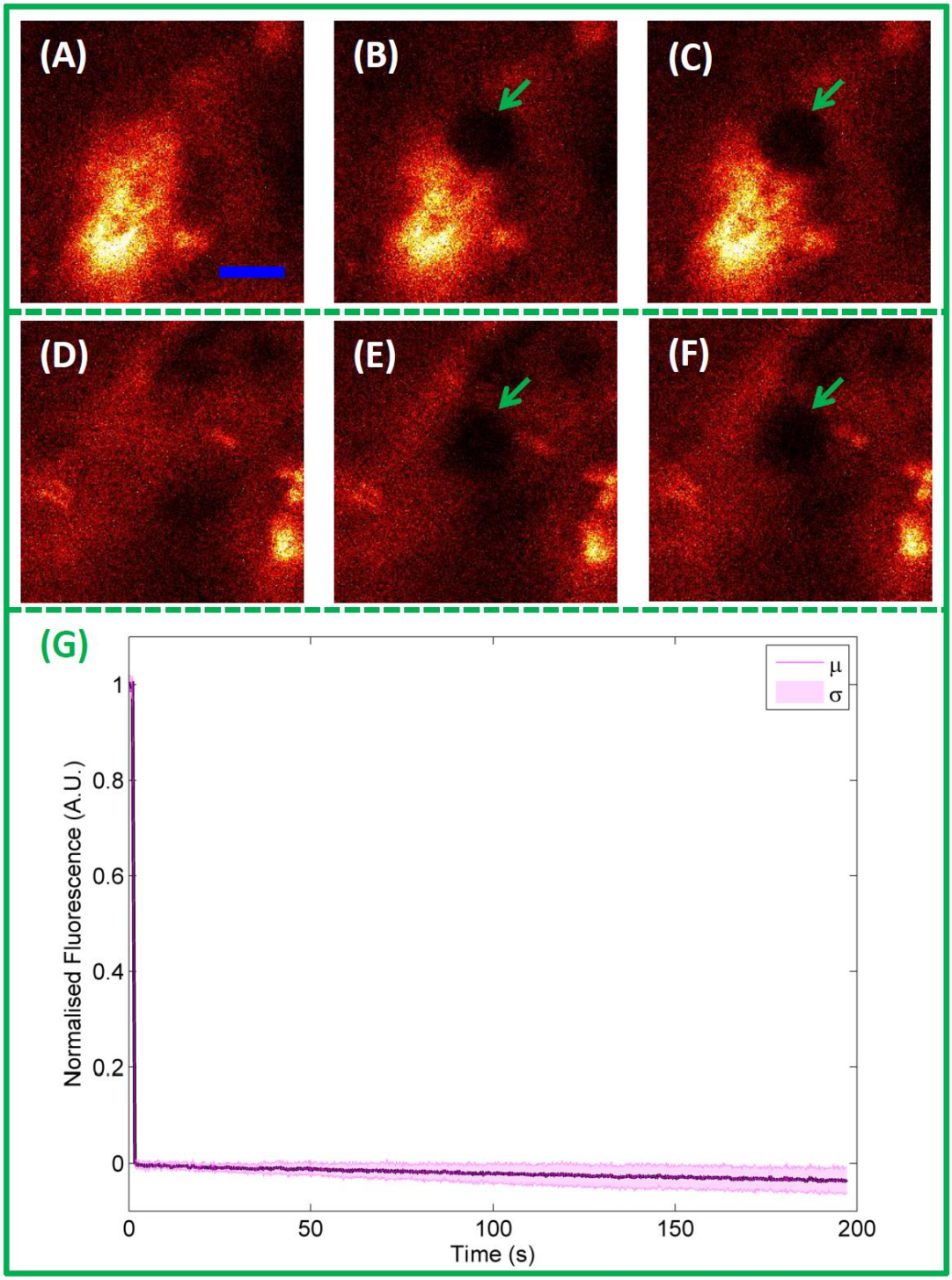


Figure 9: Diffusion of Halo-FGF10 in Rama 27 fibroblast pericellular matrix. Fixed Rama 27 fibroblasts were used to provide a pericellular matrix for Halo-FGF10 binding. A 5 μm radius disk area on the cell was bleached by full power laser to measure the recovery of the fluorescence in the bleached area. The fluorescence intensity of the bleached area was extracted to detect the diffusion of TMR-Halo-FGF10 in the pericellular matrix. **(A and D):** TMR-Halo-FGF10 labelled cells (two

areas with different binding intensities) before bleaching; **(B and E):** Same areas as (A and D), but following the bleaching of a 5 µm radius disk; **(C and F):** The partial recovery of fluorescence in the bleached area 196 s after bleaching; **(G):** The normalised fluorescence intensities of Halo-FGF10 in the bleached area were plotted against time (average of 10 measurements). Size of the scale bar: 5 µm.

Since the distribution of Halo-FGF10 in Rama 27 fibroblast pericellular matrix was very heterogeneous, FRAP experiments were conducted to determine the diffusion of Halo-FGF10 in both areas of high (Figs 9 A-C) and lower binding (Figs 9 D-F). As for Halo-FGF1, Halo-FGF2 and Halo-FGF6, a small area of the cells was bleached and the fluorescence recovery was measured over the following 196 s (Figs 9 A-C and D-F). Compared to the image acquired immediately after bleaching, there was no obvious recovery of fluorescence after 196s (Figs 9 B-C and E-F). The averaged fluorescence recovery curve demonstrates that the TMR-Halo-FGF10 in the bleaching area did not exchange appreciably with the TMR-Halo-FGF10 outside the bleached area (Fig. 9 G). These data suggest that FGF10 does not dissociate readily from the heparan sulfate, chondroitin sulfate and dermatan sulfate chains it is bound to. Interestingly, the thermal shift assay used to identify its selectivity for sulfation patterns with a library of chemically modified heparins shows that rather than equilibrating between bound and unbound forms, FGF10 appears to partition into two populations (FGF10 and FGF10 bound to heparin) (Li et al., 2015). This is consistent with a very slow dissociation of FGF10 from heparin, since faster dissociation would enable exchange of FGF10 molecules on the heparin and so an averaging of the measured thermal stability of bound and unbound species. Work in two development models where FGF10 has a role in epithelial morphogenesis, in lung and salivary gland morphogenesis, also indicates that FGF10 bound to glycosaminoglycans does not readily dissociate and that FGF10 diffusion requires either suboptimal binding structures or the action of heparanase (Izvolsky et al., 2003, Patel et al., 2007).

The substantial differences in diffusion observed between Halo-FGF1, Halo-FGF2 and Halo-FGF6 may be a consequence of differences in the number and spatial organisation of their respective binding sites on heparan sulfate chains. Alternatively, the much greater level of binding of Halo-FGF1 may reduce its mobility, due to a

lower availability of free binding sites for Halo-FGF1 to move to in the pericellular matrix. To distinguish between these possibilities a lower concentration of Halo-FGF1 was used to measure its diffusion.

Effect of changing the concentration of Halo-FGF1

The level of bound TMR-Halo-FGF1 was changed by halving the concentration of Halo-FGF1 added to fixed Rama 27, which reduced the fluorescence intensity to levels similar to that observed with 2 nM TMR-Halo-FGF2 (Fig 10 A). However, at the lower level of binding of Halo-FGF1, the recovery of fluorescence following bleaching was similar to that observed with 2 nM Halo-FGF1. Thus, the half recovery time for 1 nM Halo-FGF1 was 45 s and only 50% of the fluorescence was recovered. Consequently, reducing the amount of Halo-FGF1 bound to the heparan sulfate in pericellular matrix by a factor of 2 had no strong effect on the diffusion speed of the Halo-FGF1 or on the relative proportions that were mobile and immobile (Figs 10 B and C). These results indicate that the slower diffusion observed with 2 nM Halo-FGF1 is unlikely to be due to the larger amount Halo-FGF1 bound to heparan sulfate in the pericellular matrix. Instead the slower diffusion of Halo-FGF1 is more likely to be due to differences in the number and spatial organisation of these binding sites, and the rate of association and disassociation of the FGF1 from them.

Thus, the diffusion measurements suggest Halo-FGF1 is less mobile in pericellular matrix than FGF2 or FGF6, but still moves though only smaller steps are accessible to it. If there was a focal source of FGFs, then FGF1 would form shorter and steeper gradients than FGF2 and FGF6 in Rama 27 fibroblast pericellular matrix.

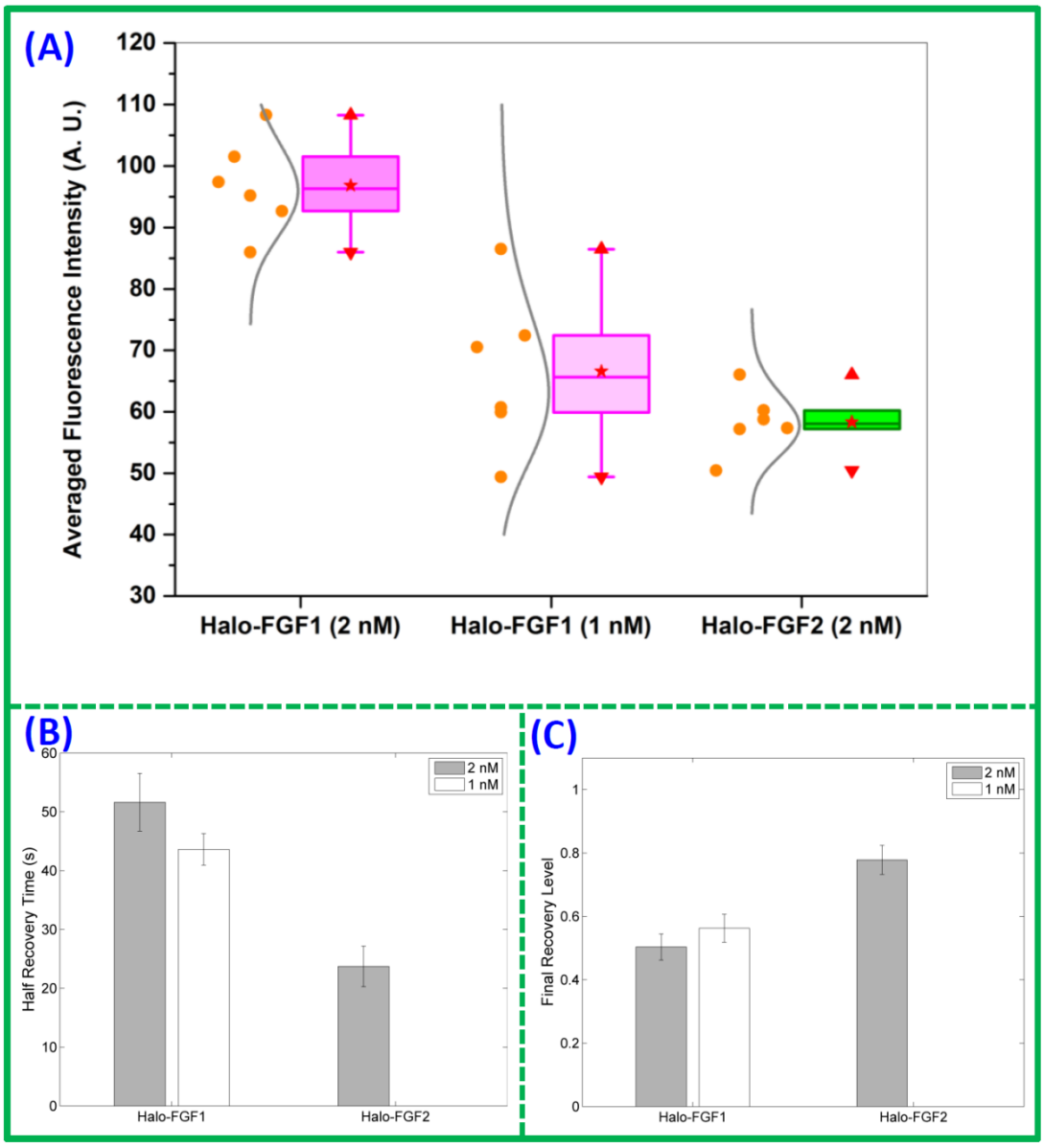


Figure 10: Effect of concentration of Halo-FGF1 and of Halo-FGF2 on their binding and diffusion in Rama 27 fibroblast pericellular matrix. TMR-Halo-FGF1 (1 nM) was used to label fixed Rama 27 fibroblasts to identify any dependence of their level of binding and their diffusion on concentration. The FRAP experiments with TMR-Halo-FGF1 (2 nM) and TMR-Halo-FGF2 (2 nM) were carried out at the same time and are those shown in Figs 7 and 8. Six FRAP experiments are included for each sample. **(A):** Binding intensities of Halo-FGF1 (1 nM and 2 nM) and Halo-FGF2 (2 nM) at the areas for FRAP measurements; **(B):** the moving speed (half recovery time) of different concentration of Halo-FGF1 and Halo-FGF2 in the matrix; **(C):** the mobile/immobile fraction (Final recovery level) of different

concentration of Halo-FGF1 and Halo-FGF2 in fixed Rama 27 fibroblasts pericellular matrix.

Binding and Movement of FGFs in extracellular matrix

The expansion of the FGF family is associated with an increase in the complexity of multicellular organisms, highlighting its importance in mediating cell communication in development and homeostasis (Itoh and Ornitz, 2004, Itoh, 2007, Itoh and Ornitz, 2011). FGFs in a subfamily are more closely related in amino acid sequence and in function than FGFs in different subfamilies. The functional relations are evidenced, for example, by the selectivity of FGFs in different subfamilies for isoforms of the FGFR (Zhang et al., 2006b, Ornitz et al., 1996), for the patterns of sulfated sugars they bind in heparan sulfate and for the number and location of heparan sulfate binding sites on the FGF (Xu et al., 2012, Li et al., 2015, Ori et al., 2009). Previous work with FGF2 demonstrated that its diffusion in pericellular matrix of Rama 27 fibroblasts was controlled by the spatial organisation of its heparan sulfate binding sites (Duchesne et al., 2012). This raises an important question, whether the diffusion of other heparan sulfate-binding effectors, with different selectivity for patterns of sulfated sugars, also possess heterogeneous networks of binding sites that control their diffusion. To tackle this question, we have used five FGFs from four different subfamilies, with well-characterised heparan sulfate binding properties (Section 4.3 and (Xu et al., 2012, Uniewicz et al., 2010)). This allows the effects of subtle differences between members of the same subfamily (FGF1 subfamily: FGF1 and FGF2) and more substantial differences between members of different subfamilies (FGF6 is in the FGF4 subfamily, FGF10 in the FGF7 subfamily and FGF20 in the FGF9 subfamily) to be measured.

There are differences between the diffusion of FGFs occurring in the experiments described here and *in vivo*. First, the Halo-FGF is bound to pericellular matrix and any unbound ligand is removed by washing. Thus, unlike *in vivo*, there is no source of diffusing ligand. Second, since the cells are fixed, receptor-mediated endocytosis cannot occur, so there is no sink to remove ligand. Therefore, the binding experiments (Figs 2-4) provide a snapshot of the distribution of binding sites on glycosaminoglycans in pericellular matrix. The FRAP experiments measure the movement of the FGF due to its dissociation and reassociation to sites on

glycosaminoglycans, without any effects of concentration gradients or cell biochemistry (membrane protein movement, membrane flow and cell movement).

There are a large number of binding sites for FGFs and other heparan sulfate binding proteins on the polysaccharide in pericellular matrix; for FGF2 in Rama 27 fibroblasts these amount to $3x10^6$ sites per cell (Duchesne et al., 2012). The five FGFs used here preferentially bind different structures in heparan sulfate (Chapter 6, Table 6.1) and, perhaps unsurprisingly, their level of binding differed considerably; FGF1 bound to the greatest extent, whereas the binding of FGF20 was undetectable, since it was within the threshold of background fluorescence (Figs 2-4 and Fig. S4). In all cases the distribution of the FGFs was heterogeneous (Figs 2-4), indicating that their binding sites are not evenly distributed in pericellular matrix. This has been shown previously for FGF2 over length scales ranging from 10 nm to several µm in the same cells (Duchesne et al., 2012, Nieves et al., 2015). The clustering of heparan sulfate proteoglycans in lipid rafts would be one mechanism that could contribute to the heterogeneous distribution of heparan sulfate binding sites (Pike, 2004, Lingwood and Simons, 2010). Other mechanisms may operate in parallel. For example, interactions of transmembrane proteoglycans, *e.g.*, syndecans (Couchman, 2003), with the cytoskeleton through their cytoplasmic domains may lead to their localisation to particular membrane microdomains.

The present data demonstrate that the heterogeneous distribution of binding sites observed previously with nanoparticle labelled FGF2 (Duchesne et al., 2012, Nieves et al., 2015) and in experiments with radiolabelled FGF2 (Chu et al., 2004) is likely to be a more general phenomenon, since it was observed here also with FGF1, FGF6 and FGF10. One interpretation is that the heparan sulfate chains possessing binding sites for a particular protein (FGFs in the present case) are at least in part differently localised in pericellular matrix, through, for example, the various clustering mechanisms discussed above. However, this interpretation is likely to be too simplistic. For heparan sulfate there are 883 extracellular proteins that bind it in the human proteome (Ori et al., 2011, Nunes, 2015). Thus, the subset of the heparan sulfate-binding proteins expressed by Rama 27 fibroblasts will have a substantial portion of their binding sites engaged with heparan sulfate. Consequently, the heparan sulfate binding sites available to a particular FGF ($3x10^6$ for FGF2 in Rama

27 cells) are likely to be less than the total possible binding sites (Duchesne et al., 2012). Moreover, these heparan sulfate-binding proteins also have very extensive networks of protein-protein interactions (Ori et al., 2011), which will influence their protein-polysaccharide interactions. One consequence of this multiplicity of interactions is that there are many free binding sites for exogenously-added proteins on heparan sulfate (*e.g.*, Figs 2-4) and there are many free binding sites on endogenous heparan sulfate-binding proteins for exogenously added polysaccharide (Castellot et al., 1985). Thus, pericellular matrix is not at equilibrium and the ingress of a heparan sulfate binding protein may perturb a wide range of interactions. Such perturbations may involve substantial changes in the 3-dimensional structure of heparan sulfate chains. For example, a number of heparan sulfate binding proteins are multivalent, that is they have more than one binding site for the polysaccharide (Haas and Culp, 1984, Battaglia et al., 1992, Sweeney et al., 1998, Migliorini et al., 2015, Xu et al., 2012, Lortat-Jacob et al., 2002, Li et al., 2015, Ori et al., 2009). A recent biophysical analysis of brushes of heparan sulfate chains demonstrated that some heparan sulfate binding cytokines and growth factors with multiple binding sites are able to cross link the chains (Migliorini et al., 2015). Since heparan sulfate - binding matrix proteins such as collagens and fibronectin have multiple binding sites for the polysaccharide, it seems reasonable that they too will in some instances cross-link heparan sulfate chains. Thus, the heparan sulfate chains in pericellular matrix are likely to be engaged in large-scale supramolecular networks, which would ultimately be responsible for the heterogeneous distribution of binding sites and through which the Halo-FGFs diffuse.

The FRAP data for the four FGFs with detectable binding show that they move differently in Rama 27 fibroblast pericellular matrix (Figs 7-10). In the case of FGF1 and FGF2, the slower movement of the former may be explained by its larger number of binding sites. Within the FGF1 subfamily, FGF1 binds to any desulfated saccharide structure of degree of polymerisation (dp) 4 or longer, whereas FGF2 requires N-sulfate and 2-O sulfate groups (Turnbull et al., 1992, Maccarana et al., 1993, Jemth et al., 2002, Uniewicz et al., 2010). Thus, even taking into account occupation of some sites by endogenous proteins, the greater promiscuity of FGF1 is likely to explain why Halo-FGF1 binds Rama 27 pericellular matrix to a greater extent than Halo-FGF2. The larger number of sites in heparan sulfate that FGF1 can

bind may also underlie its more restricted mobility; a greater density of binding sites would reduce the distance the protein can travel in a given time, since the likelihood of rebinding will be greater. Indeed, binding site density and clustering have been shown to prevent effective dissociation of heparan sulfate binding proteins such as FGF2 from pericellular matrix and is likely to alter the distance a protein can travel within pericellular matrix before re-binding (Dowd et al., 1999). The differences in movement of the other FGFs would then similarly reflect their selectivity for binding structures in heparan sulfate and how the available binding structures are presented. In the extreme, as seen with FGF10, the FGF does not diffuse appreciably over the time of the FRAP measurement. In such instances, the movement of the heparan sulfate-binding protein would require additional mechanisms. This could be provided by heparanase, an extracellular β glucuronidase, which cleaves heparan sulfate chains in their unmodified and transition domains. This would release cargoes of S-domains and bound protein, as shown for FGF2 in a skin wound healing model (Kato et al., 1998). Indeed heparanase has been shown to be important for the stimulation of ductal morphogenesis by FGF10 in salivary gland (Patel et al., 2007).

Conclusion:

The selectivity of FGFs for different binding structures in glycosaminoglycans provides a means to probe the distribution of these binding sites in Rama 27 cell pericellular matrix and to determine the effect this has on the diffusion of the FGFs. The results show that protein binding sites in heparan sulfate (and chondroitin sulfate/dermatan sulfate for FGF10) of pericellular matrix are not randomly distributed. A number of different mechanisms are likely to regulate the distribution of these binding sites, including the biosynthesis of the heparan sulfate chains, the localisation of core proteins in membrane microdomains and the interactions of the polysaccharide chains with endogenous heparan sulfate binding proteins. The high multiplicity of interactions, between both proteins and polysaccharide and between the polysaccharide-binding proteins themselves (Ori et al., 2008, Ori et al., 2011, Xu and Esko, 2014) are likely to produce a dynamic network of interlinked molecules. This would then be responsible for the long-range (supramolecular) structure of the pericellular matrix, which determines its spatial binding capabilities for individual proteins. Such a structure would be sensitive to perturbations, such as the ingress of a

heparan sulfate binding protein from a neighbouring cell (in the same or different tissue compartment) and clearly can control the diffusion of such effector proteins. Supramolecular structure in extracellular matrix has been clearly shown in cartilage (Heinegard and Saxne, 2011), where there are also definitive structural and functional differences between the pericellular matrix of chondrocytes, and the territorial and inter-territorial matrices that are more distant from the cells. Thus, although extracellular matrix in cartilage is specialised, in other tissues an analogous situation may exist, where pericellular, extracellular and basement membrane matrices may exhibit different types of supramolecular structure and consequently has different functions.

5.3 Supplementary results

5.3.1 Supplementary results for the manuscript in Section 5.2

The bright channel images in Figs S1, S2, S3 and S4 were acquired to show the locations and profiles of the cell. The edges of the cells with low fluorescence intensity were highlighted as described in the methods. Fluorescence and brightfield channels of Halo-FGF20 (Figs. S4 A and C) and HaloTag (Figs. S4 B and D) show the binding of Halo-FGF20 to the Rama 27 fibroblast pericellular matrix. Since Halo-FGF20 exhibited weak non-specific binding to the culture dish (Fig. S4 A), any binding of Halo-FGF20 was within this background signal.

The fluorescence intensities in a same area were quantified for Halo-FGF1, Halo-FGF2 and Halo-FGF6 over 4.5 min (Fig. S5). No change of the fluorescence intensities for these three Halo-FGFs was detected, indicating that the FGFs are effectively trapped in pericellular matrix over this time and do not dissociated appreciably into the bulk culture medium. The fluorescence intensities of reference areas during the FRAP experiments were also quantified. The FGF trapping results (Fig. S5) indicate that the changes of fluorescence intensities of reference areas during FRAP experiments (Fig. S6) were caused by photobleaching and exchange of the bleached TMR-Halo-FGFs from the bleaching area, rather than dissociation of

Halo-FGF into the bulk culture medium. The imaging videos are examples of FRAP experiments for Halo-FGF1, Halo-FGF2 and Halo-FGF6.

Figure S1: The brightfield images corresponding to the images in Fig 2. The images A, B, C, D, E, F, G, H and I are corresponding to images A, B, C, D, E, F, G, H and I in figure 2.

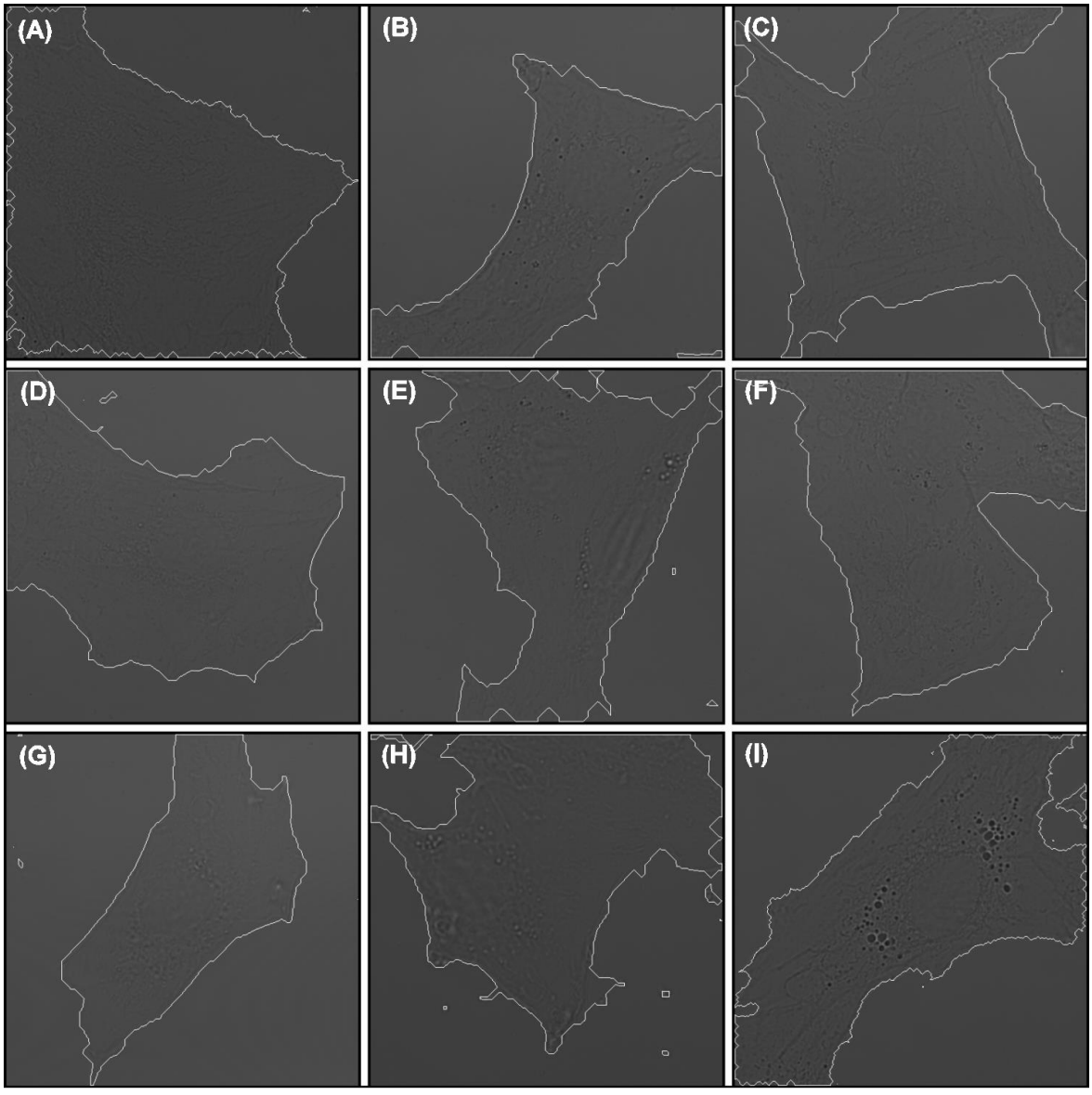


Figure S2: The brightfield images corresponding to the images in Fig 3. The images A, B, C, D, E and F are corresponding to images A, B, C, D, E and F in figure 3.

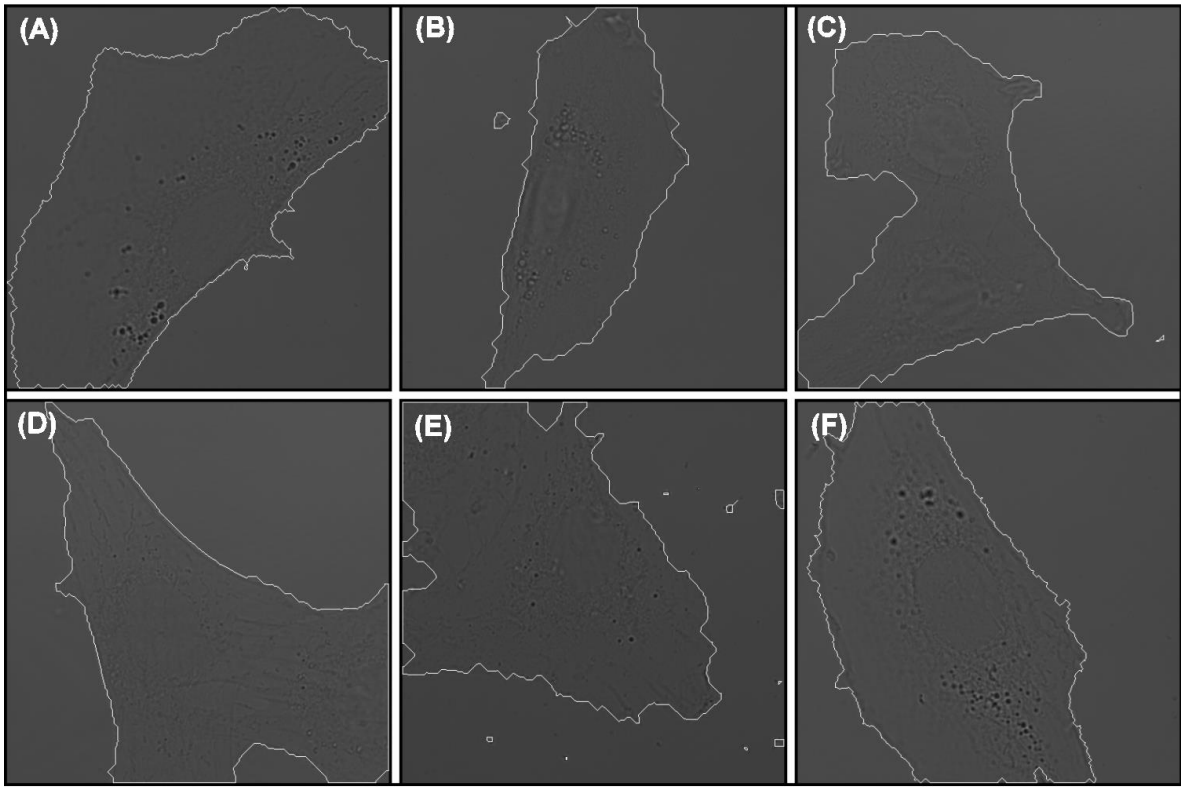


Figure S3: The brightfield images corresponding to the images in Fig 4. The images A, B, C, D, E, F, G, H and I are corresponding to images A, B, C, D, E, F, G, H and I in figure 4.

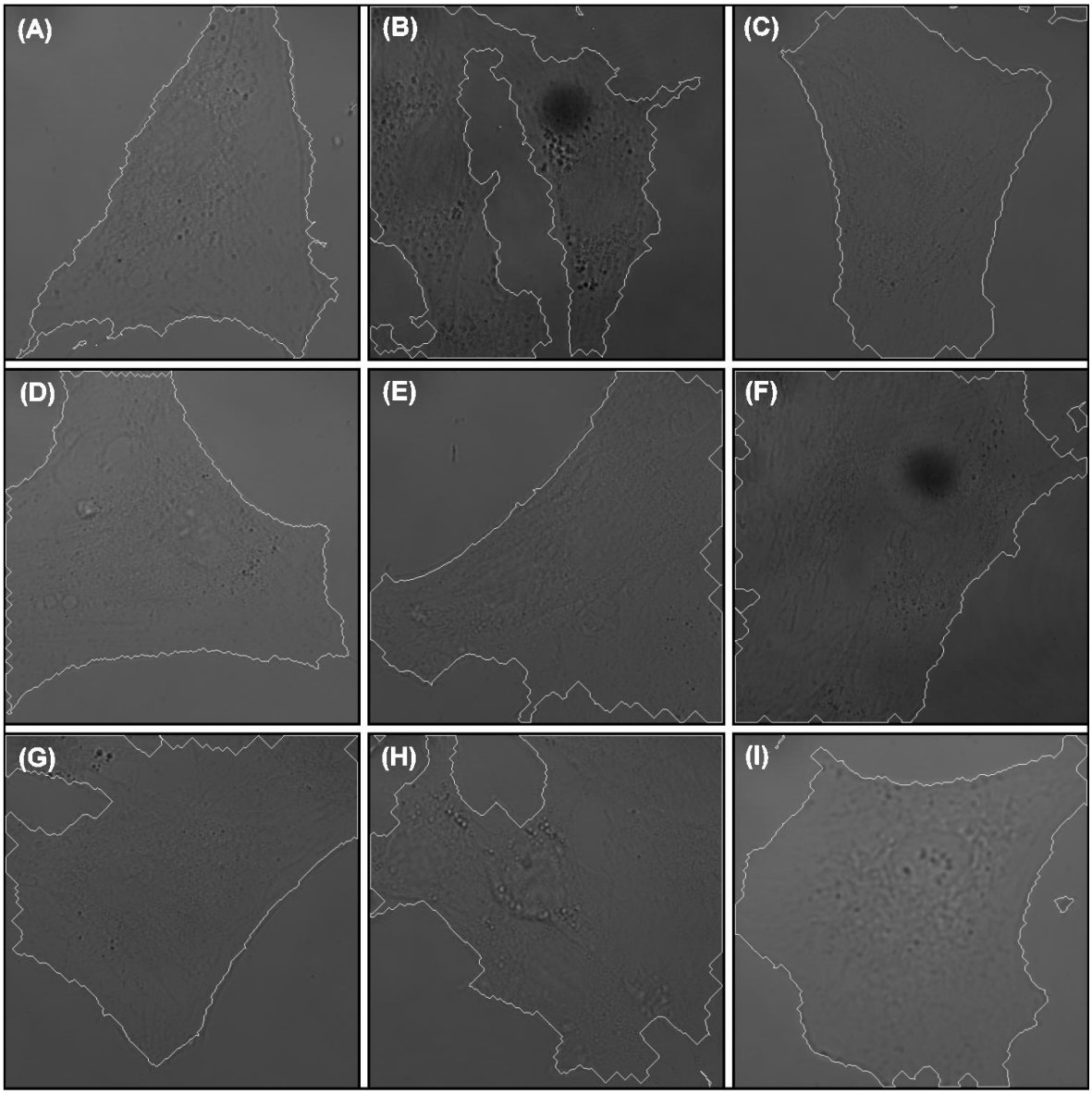


Figure S4: Binding of Halo-FGF20 and HaloTag to Rama 27 fibroblasts: (A): Rama 27 fibroblasts were incubated with 2 nM TMR-Halo-FGF20; **(B):** Rama 27 fibroblasts were incubated with 2 nM TMR-HaloTag; **(C-D):** The brightfield images channel for (A) and (B). Image size: 67.3 µm X 67.3 µm.

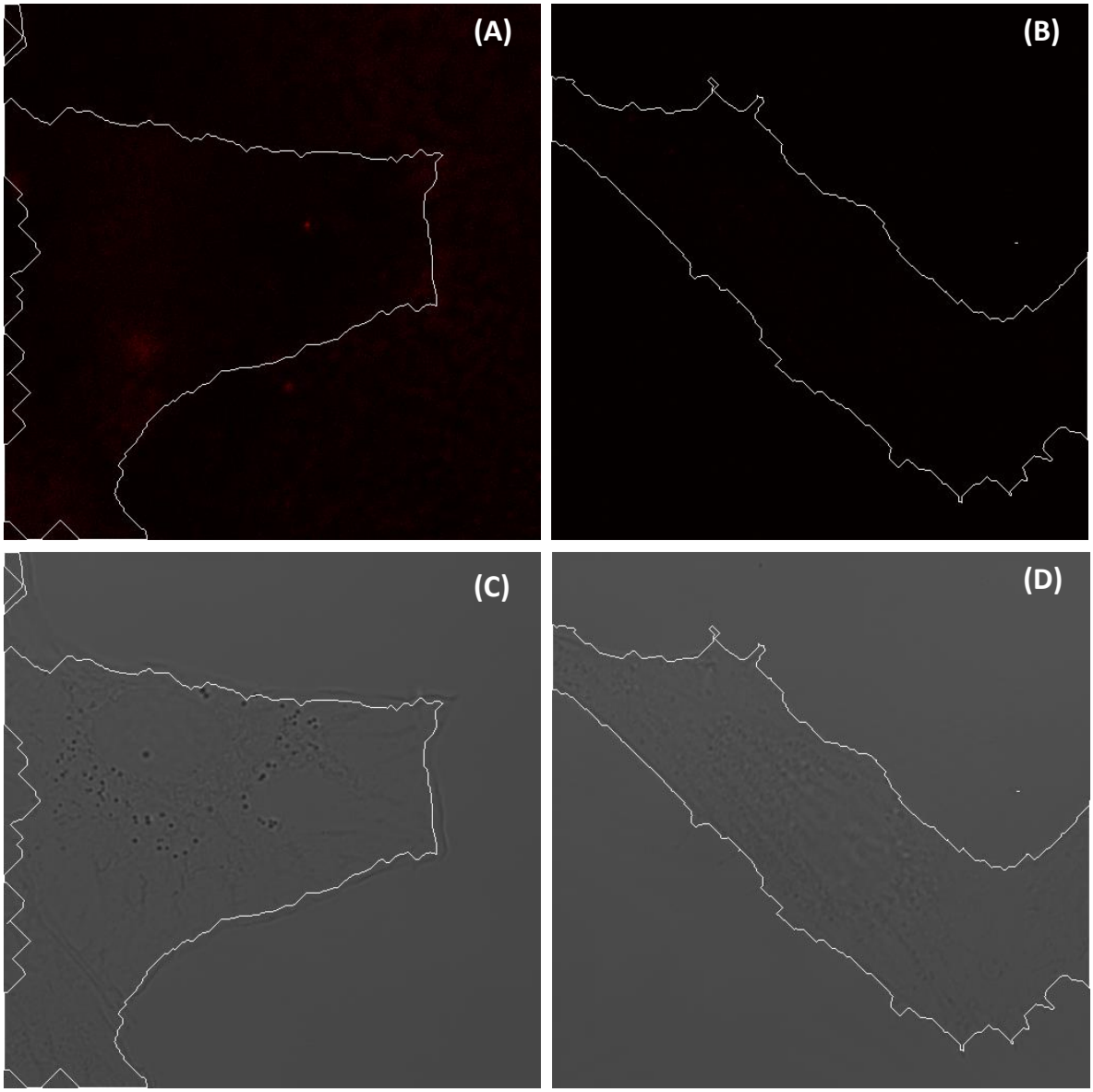


Figure S5: Dissociation of Halo-FGF1, Halo-FGF2 and Halo-FGF6 from Rama 27 pericellular matrix into the bulk culture medium. Rama 27 fibroblasts were incubated with 2 nM TMR-Halo-FGF1, TMR-Halo-FGF2 and TMR-Halo-FGF6. Ten images at 30 s interval were acquired for each area (six different areas for each Halo-FGF) to measure the dissociation of Halo-FGFs from the pericellular matrix into the bulk culture medium. The averaged fluorescence intensity was quantified as described in materials and methods. The fluorescence intensity lines above the dash line are the fluorescence intensities of Halo-FGF1. Halo-FGF2 is between the dash line and dotted line and Halo-FGF6 is below the dotted line.

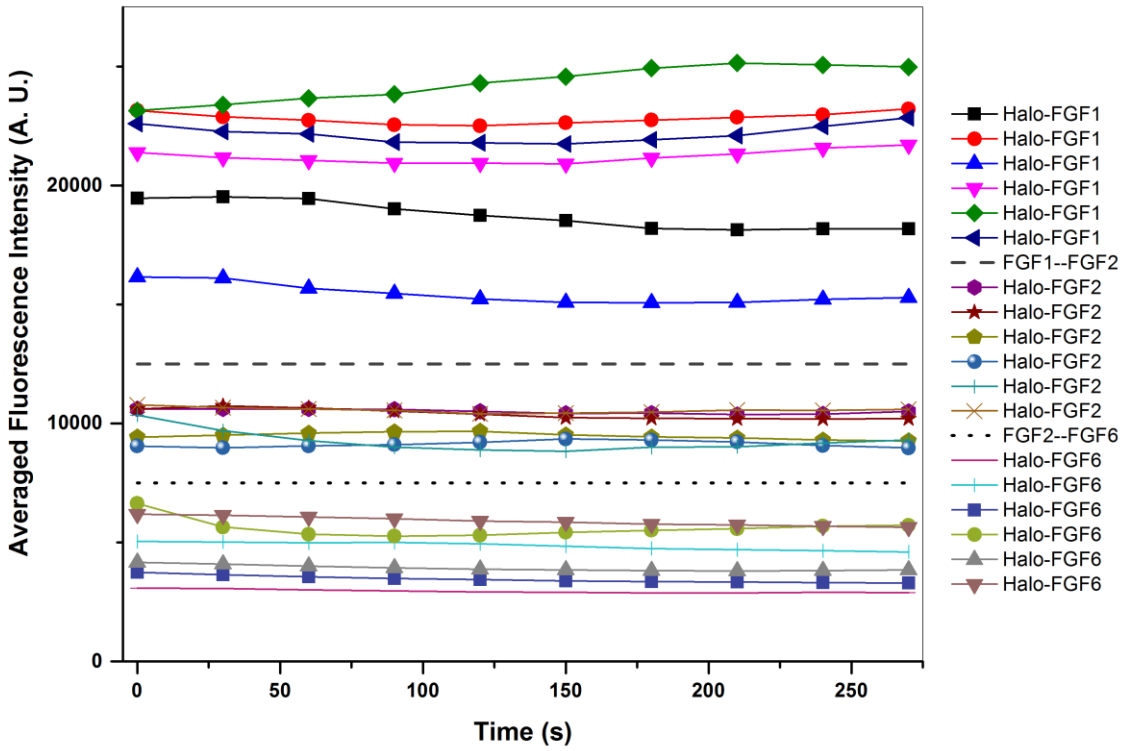
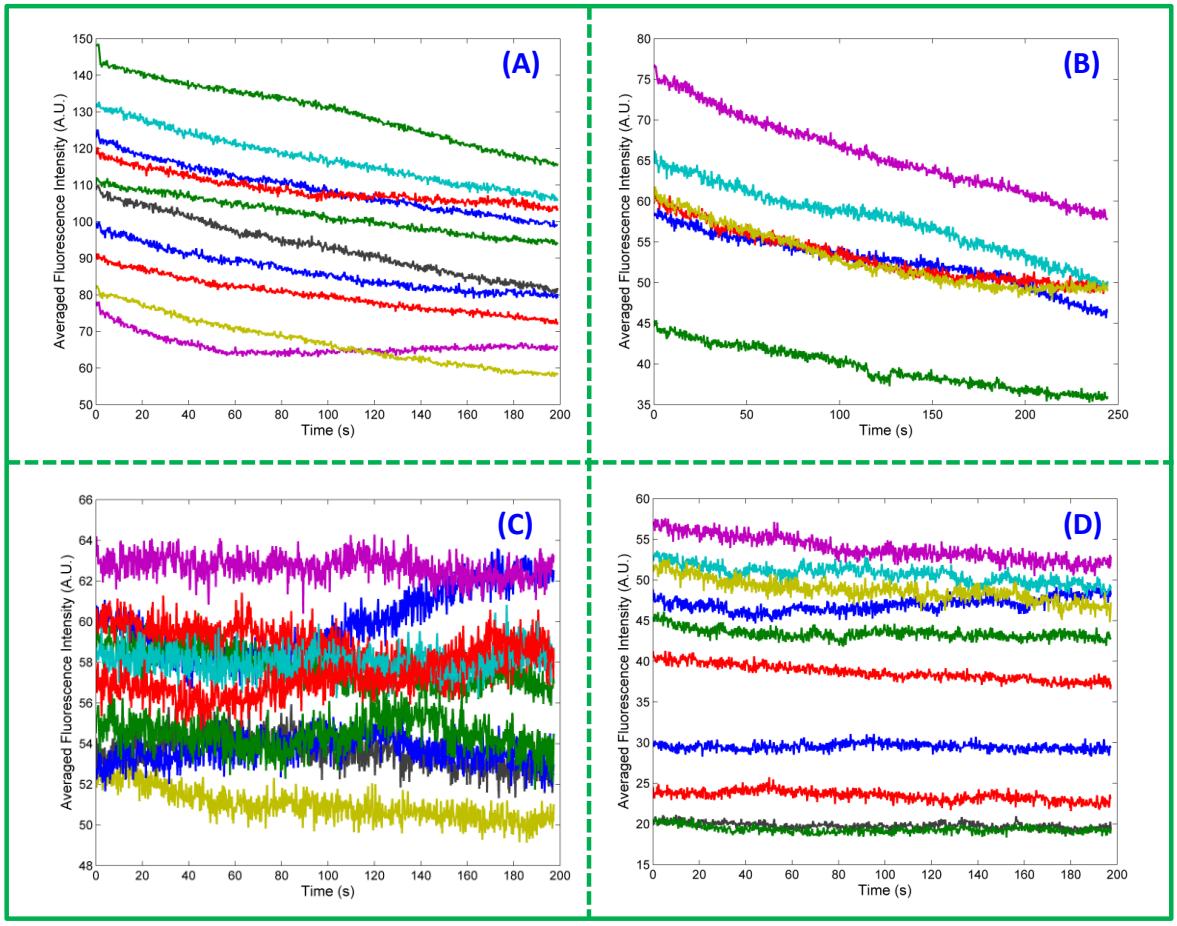


Figure S6: Photobleaching of TMR-Halo-FGFs during FRAP experiments. The fluorescence intensity of the reference area was extracted to determine the photobleaching of TMR-Halo-FGFs during the FRAP experiments. **(A):** Averaged fluorescence intensity of the reference area labelled with 2 nM Halo-FGF1; **(B):** Averaged fluorescence intensity of the reference area labelled with 1 nM Halo-FGF1; **(C):** Averaged fluorescence intensity of the reference area labelled with 2 nM Halo-FGF2; **(D):** Averaged fluorescence intensity of the reference area labelled with 2 nM Halo-FGF6.



Video S7-S10: FRAP movies for 2 nM Halo-FGF1 (S7), 1 nM Halo-FGF1 (S8), 2 nM Halo-FGF2 (S9) and 2 nM Halo-FGF6 (S10). Each video consists of 1000 images acquired over 197 s.

5.3.2 Nonspecific binding of Halo-FGFs to the glass dish

A confounding issue in imaging is the non-specific binding of proteins and probes to the glass dish or cells. This is often blocked by incubating with BSA. However, some FGFs were found to bind to the glass or BSA blocked dish, even when BSA in the binding buffer was used to compete for the nonspecific binding. When Halo-FGF7 was used to label Rama 27 cells, as described for Halo-FGF2, more fluorescence was found on the cell free area than on the cell membrane. This indicates that the non-specific binding signal was stronger than the potential specific one.

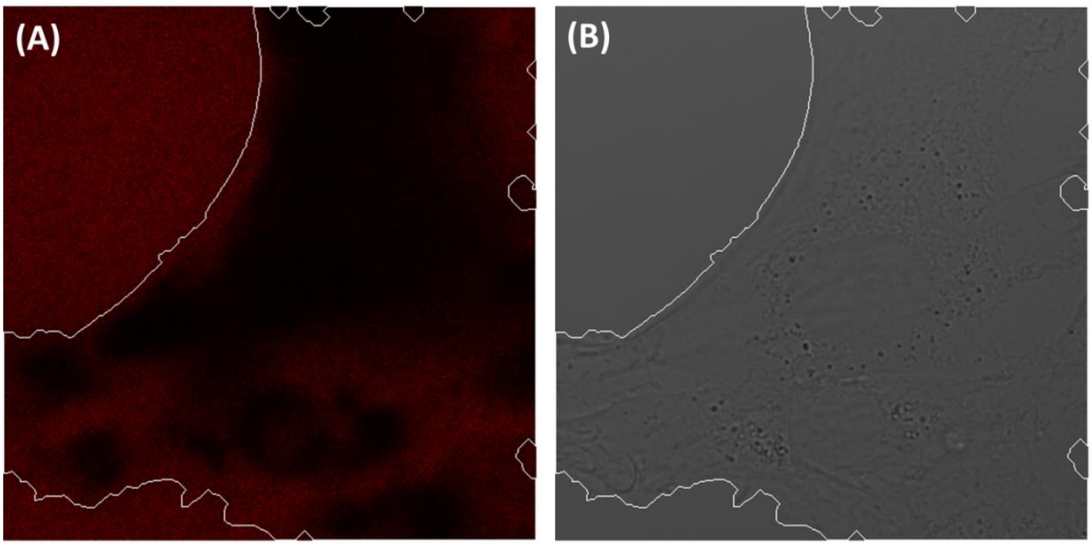


Figure 5.1 Using TMR-Halo-FGF7 to label fixed Rama 27 fibroblasts. (A): Rama 27 fibroblasts were incubated with 2 nM TMR-Halo-FGF7; **(B):** The bright field channel for (A). Image size: 67.3 µm X 67.3 µm.

5.3.3 Cross-binding competition of FGFs to Rama 27 fibroblast pericellular matrix

TMR-Halo-FGF1 and TMR-Halo-FGF2 (both 2 nM) were competed with seven different concentrations of FGF1 and FGF2. Thus, TMR-Halo-FGF1 (2 nM) was

used to label fixed Rama 27 fibroblasts in the presence of competing FGF1 (8 μM, 2 μM and 0.4 μM) or of FGF2 (8 μM, 2 μM, 0.4 μM and 0.1 μM) in a 37°C incubator for 30 min. A similar experiment was conducted with TMR-FGF2. The excess FGF was removed by three washes with PBS. The images were collected as before (Section 5.2), to determine the changes of binding intensities of TMR-Halo-FGF1 and TMR-Halo-FGF2.

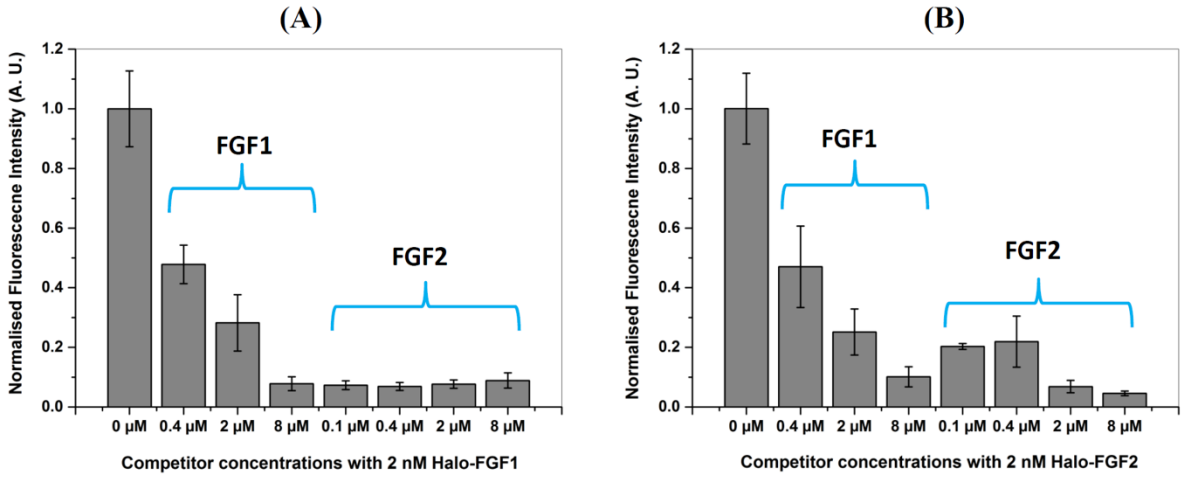


Figure 5.2 Competition of Halo-FGF1 and Halo-FGF2 binding to Rama 27 fibroblasts by FGF1 and by FGF2. The labelling of 2 nM Halo-FGF1 and 2 nM Halo-FGF1 was competed with different concentrations of FGF1 and FGF2, as indicated in the figure. **(A):** Binding intensity of 2 nM Halo-FGF1 in the presence of competing unlabelled FGF1 and FGF2. **(B):** Binding intensity of 2 nM Halo-FGF2 in the presence of competing unlabelled FGF1 and FGF2.

When 0.4 μM FGF1 was used to compete for the binding of Halo-FGF1, the binding of TMR-Halo-FGF1 decreased to half that of the control sample (no competitor) and it required 8 μM FGF1 to reduce binding by over 90% (Fig. 5.2). In contrast, 0.1 μM

FGF2 could prevent most of the binding of TMR-Halo-FGF1. At first sight this is somewhat puzzling. FGF1 has more binding sites in Rama 27 pericellular matrix (Figs 2 A and 3 A, Section 5.2), which is consistent with FGF1 possessing binding structures in HS than FGF2, since it binds heparin structures containing two of N-sulfate, 6-O-sulfate or 2-O-sulfate, whereas FGF2 only binds desulfated structures containing both N-sulfate and 2-O-sulfate (Turnbull et al., 1992, Maccarana et al., 1993, Jemth et al., 2002, Uniewicz et al., 2010). Therefore, it is expected that there are binding sites occupied by FGF1, which FGF2 cannot compete for. When the experiment was reversed, so the FGF2 was labelled, the inclusion of unlabelled FGF1 could be interpreted as a simple competitor. With 0.4 μM unlabelled FGF1, the fluorescence of Halo-FGF2 was decreased by around half and with 8 μM FGF1, it was decreased by about 90%. In contrast, 0.1 μM FGF2 could reduce the binding of Halo-FGF2 by 80% and 2 μM FGF2 reduced the binding by about 90%. These competition results indicated that FGF2 is a better competitor than FGF1 for FGF1 and FGF2 binding sites.

These data show that competition by FGF1 for FGF2 binding sites (and *vice-versa*) in Rama 27 pericellular HS does occur. The stronger competition by FGF2 for FGF1 and FGF2 binding sites could be explained by the absence of FGF1 binding sites containing desulfated saccharide sequences with 6-O sulfate. However, this would contradict the much higher level of binding of FGF1 than of FGF2 to this pericellular matrix. An alternative explanation is that the experiment is probing mechanisms rather than simple mass action-driven displacement. It is established that FGF2 is able to crosslink HS chains, at least *in vitro* (Migliorini et al., 2015), though whether FGF1 can do this is not known. In any event, if FGF2 were crosslinking HS chains,

this would cause substantial conformational change in the polysaccharide and may as a consequence reduce the affinity of FGF1 for binding sites unoccupied by the competing FGF2. In this scenario, FGF2 exerts a competitive effect directly by displacement of bound FGF1 and indirectly, by changing the conformation of HS chains.

5.4 Matlab programs for imaging data analysis

Using DrawCellEdges (a Matlab user interface) to highlight the cell edges and to quantify fluorescence intensity of individual cells

Program introduction

The DrawCellEdges program (Fig. 5.3) was developed to draw the cell edges using either the fluorescence channel or the bright field channel. The cell edges are also drawn on another channel to highlight the cell areas. The average fluorescence intensity of the cell areas is automatically quantified and saved to an excel file for further analysis.

The program was developed by following a published Matlab code (MathWorks, 2014), and some important parameters in the 'Draw Cell Edges' panel can be modified to adapt to the user's input image. Multiple separate images (single or multiple channels for each image) or an image stack containing many images can be analysed by running the program once. The details for using the program are introduced below.

File table:

DrawCellEdges.m contains the main code for analysing the imaging data.

DrawCellEdges.fig is a figure window, which contains the user interface, and all the functions in the interface are connected to the main code to do the mathematical analysis.

User instruction:

1. **Run the program.** Open DrawCellEdges.m in MATLAB R2014a installed on a windows computer and run this program. The user interface (Fig. 1) will be displayed on the screen.
2. **Load the first image (Fig. 5.3 'Load First Image' panel).** If multiple separate images are to be analysed, the image files should be named as filename**0n**.lsm (n is a number from 1 (the first image) to n (number of the last image)). There is no special name requirement for an image stack. The fluorescence channel, bright field channel and the image bit-depth should be specified in the text box. Then, the first image or the image stack can be loaded by pushing 'Load' button and selecting the image file in the desired directory.

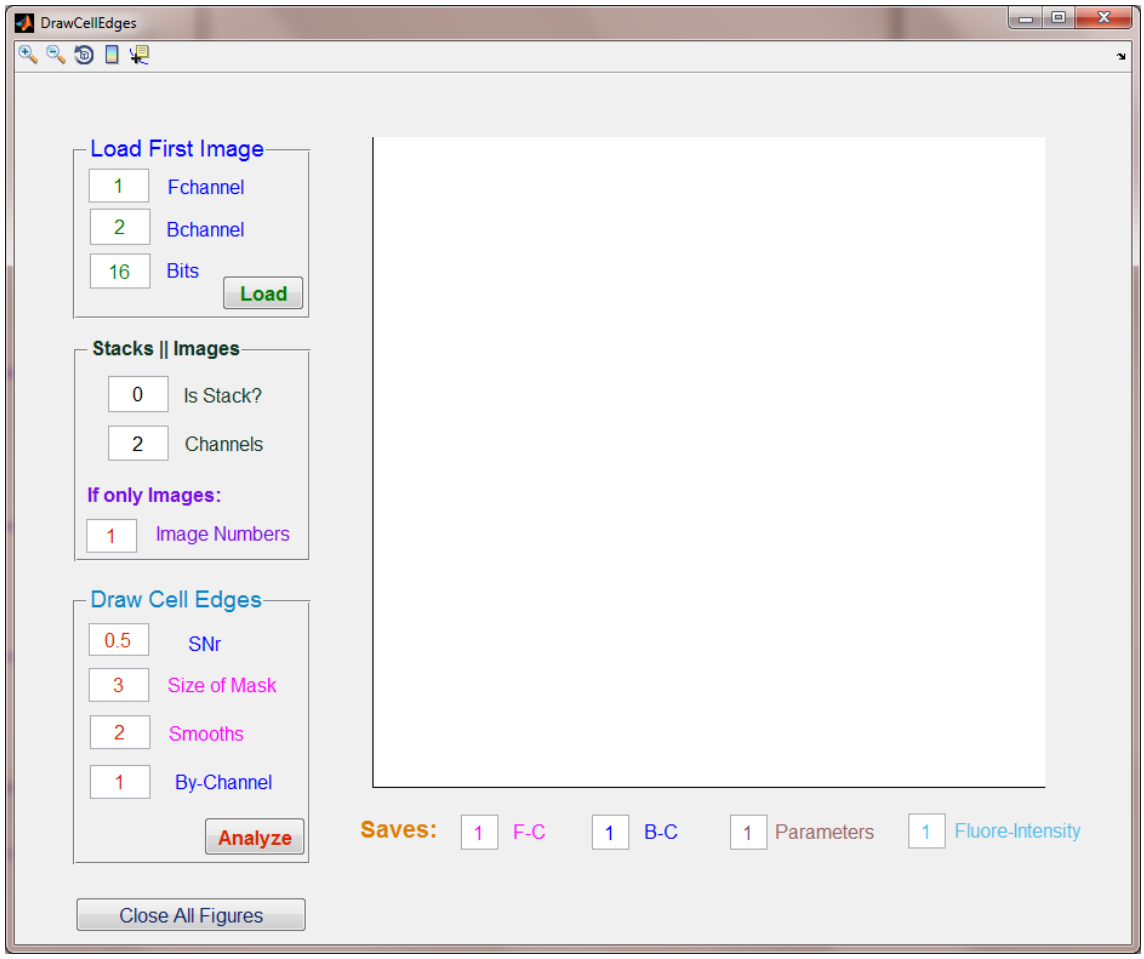


Figure 5.3 User interface of DrawCellEdges program. The user interface contains the parameters panels on the left side and image display box on the right side. The save functions are under the image display box.

3. **Select the image stack or multiple images (Fig. 5.3 'Stacks || Images' panel).** If multiple images are to be analysed, the text box for 'Is stack' should be 0, which means no. The total number of images with same filename ('Image Numbers' text box) should be specified to analyse multiple images in a single run. If an image stack is to be analysed, the text box should be 1, which means yes. The total number of images in the stack ('Image Numbers' text box) should also be specified if the images in the stack are to be run once as a single batch. The number of total channels is required as well to logically analyse the desired image.
4. **Load parameters and analyse the cell edge (Fig. 5.3 'Draw Cell Edges' panel).** The loaded image/images can be analysed with the default parameters. If the acquired results are not fit for purpose, the following parameters (MathWorks, 2014) can be modified to re-analyse the images.

The 'SNr' is a threshold parameter (0~1) used to separate the cell from background. The recommended range for this parameter is 0.35~0.65. The 'Size of Mask' (generally around 3) is used to dilate the image. Since the detected cell areas generally contain some gaps, the image should be dilated to fill the gaps. The 'Size of Mask' is based on the size of gaps. 'Smooths' is used to remove the small noisy areas, which are detected outside of cell areas. This positive integer number is the number of cycles for smoothing the image. The recommended number is 2~6 cycles.

5. **Saves.** The analysed images (fluorescence channel and bright filed channel) with highlighted cell edges are automatically saved in 'analysed' folder in the same directory with the loaded images. The parameters for image analysis and the averaged fluorescence intensities of the cell areas are saved in the same folder as well. The save function can be turned off by changing the corresponding parameters to 0.

The analysed images will be displayed on the screen and all of the figures can be closed by pushing 'Close All Figures' button.

Using ImagingDataAnalyzerForFRAP (a Matlab user interface) to analyse fluorescence recovery after photobleaching data

Program introduction

The ImagingDataAnalyzerForFRAP program was developed to efficiently analyse fluorescence recovery after photobleaching (FRAP) data. Three main functions are included in the program: fluorescence recovery curve analysis, imaging data viewing and radial profile analysis (Fig. 5.4).

The fluorescence intensity of the bleached area is analysed in box 1 (Fig. 5.4: fluorescence recovery curve analysis) with the method described in materials and methods (manuscript). The half recovery time and final recovery level are also extracted (Fig. 5.5). The raw FRAP imaging data can be loaded and viewed in box 2 (Fig. 5.4: loading and viewing imaging data). The radial profile of the bleached area and surround is analysed in box 3 (Fig. 5.4: radial profile analysis) to spatially investigate the fluorescence recovery. By default, the frame numbers (or image numbers) corresponding to before bleaching, immediately after bleaching and from 10% to 100% final recovery level in steps of 10% were extracted from the analysed fluorescence recovery curve (Fig. 5.5). The radial profiles of these twelve frames are extracted. More details of the files included in this programs and a user instruction are described below.

File table:

ImagingDataAnalyzerForFRAP.m (ImagingDataAnalyzerForFRAP.fig) is the main script, which contains the program interface (Fig. 5.4) and the code for fluorescence recovery curve analysis. The view function of raw imaging data (.lsm format) is included in this script as well.

rscan.m is a published Matlab code for radial profile extraction (Chattrapiban, 2007). A profile example is presented in figure 2. The profile of one image was extracted from the bleaching centre to the circle (Fig. 5.6).

MoRscan.m is a linker between the user interface and rscan.m. It passes the parameters to rscan.m and returns the collected radial profile data back to the main script.

shadedErrorBar.m is a published Matlab code (Campbell, 2010), which is used to add a shaded error bar to the plotted curve.

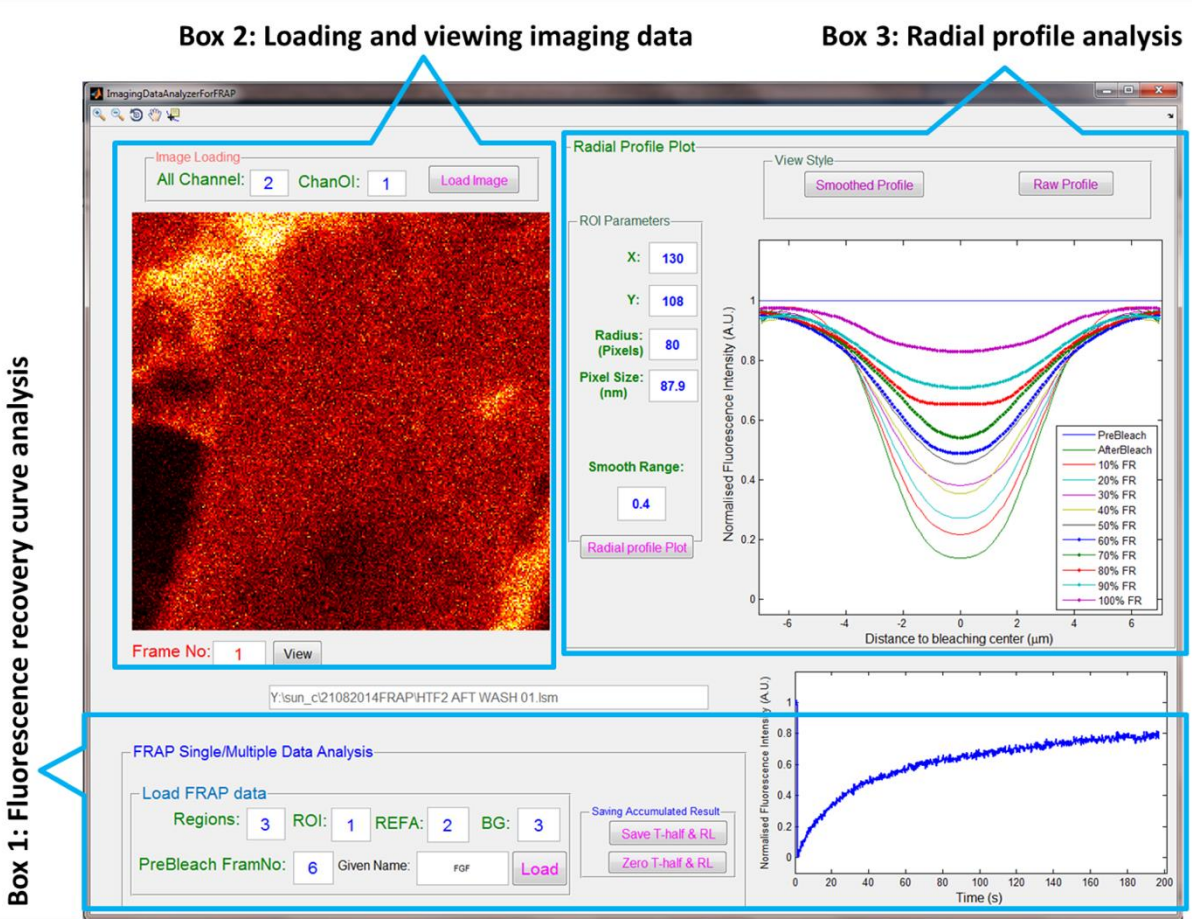


Figure 5.4 User interface of ImagingDataAnalyzerForFRAP. Three main functions are shown in the interface. Box 1 is the fluorescence recovery curve analysis panel. Box 2 is used to view the images from the FRAP experiments. Box 3 combines the desired image numbers from box 1 and the images from box 2 to extract the radial profiles. The default parameters are the parameter used for the Halo-FGFs study.

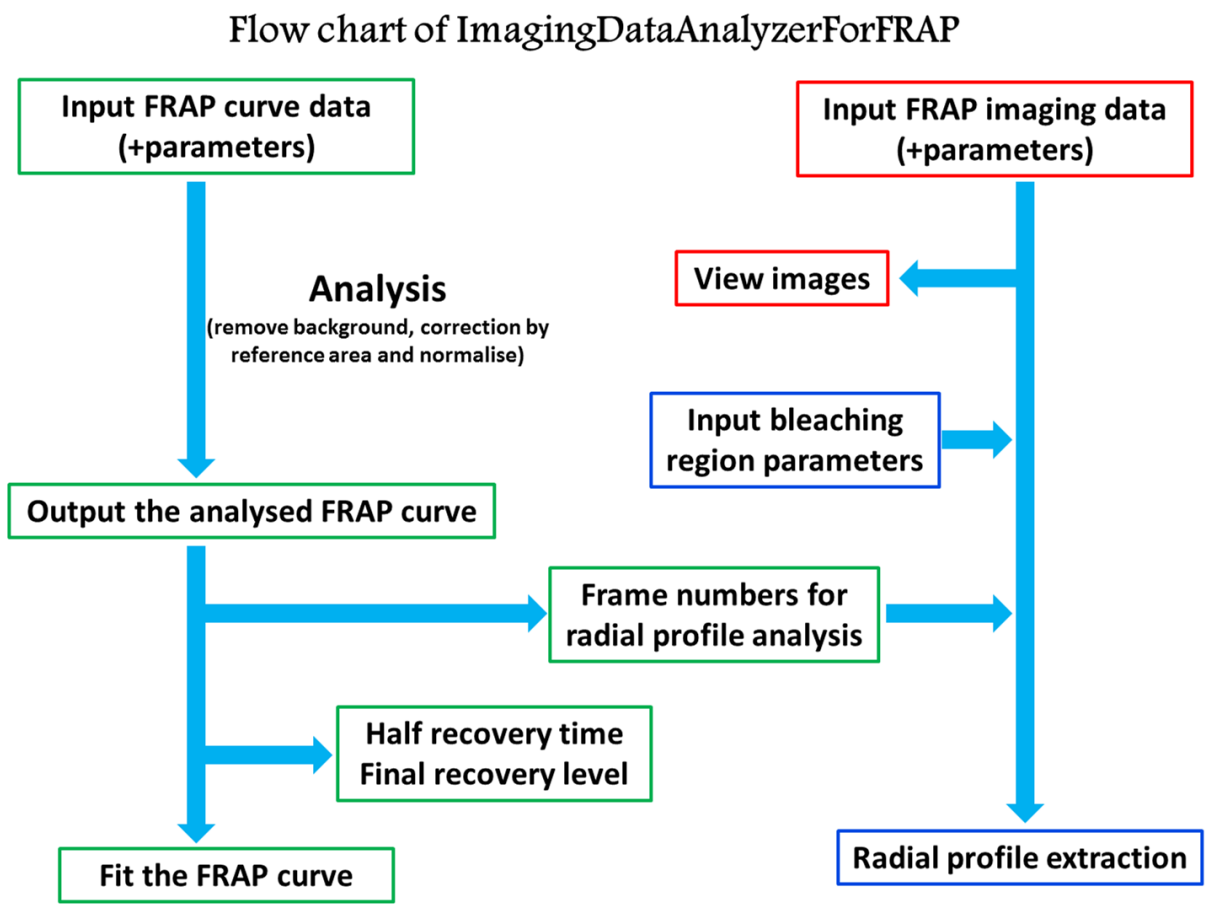


Figure 5.5: Flow chart of ImagingDataAnalyzerForFRAP. The diagram shows how the program works and how the different functions communicate with each other.

User instruction:

1 Run the program. Download shadedErrorBar.m (Campbell, 2010) and rscan.m (Chattrapiban, 2007) from the webpage link in the reference list and add these two files to the folder containing the program. Open ImagingDataAnalyzerForFRAP.m in MATLAB R2014a installed on a windows computer and run this program. The user interface (Fig. 5.4) will be displayed on the screen and be ready to use.

2 Fluorescence recovery curve analysis

The fluorescence recovery curve is not fitted, but the FRAP curve fitting was described in another FRAP software paper (Rapsomaniki et al., 2012). As the half recovery time and final recovery level are extracted, a simulation curve is generated to test whether the fluorescence recovery curve can be fitted to the model:

Model: $I = I_f - I_f * e^{(-\ln(2) * t/t_{1/2})}$

I: Fluorescence intensity at time t; I_f: Final fluorescence intensity; t: time value; $t_{1/2}$: Half recovery time.

The parameters in box 1 (Fig. 5.4) are based on the fluorescence recovery input data. In general, the fluorescence intensities of three regions are included in the input data: fluorescence intensity of bleached area (ROI), fluorescence intensity of non-bleached reference area (REFA) and fluorescence intensity of background (BG). The number of regions and column position of the three regions in the file are required to be specified. The time values are always in first column of the input data, but it is not counted for the column position of the three regions. The number of images collected before bleaching should be input into the 'PreBleach FramNo' text box. 'Given Name' will give a name to the analysed half recovery time and final recovery level, but it is not required if the half recovery time and final recovery level will not be saved in an excel file by 'Save T-half & RL' button ('Zero T-half & RL' function is used to remove from memory the half recovery time and final recovery level).

The 'Load' button allows a file to be chosen (excel or txt) for analysis. A file containing both a single fluorescence recovery curve and multiple fluorescence recovery curves can be analysed. The analysed fluorescence recovery curve and the fluorescence intensity of the reference area will be presented in separate figures.

3 Load imaging data

The parameters for the image channel should be checked before loading the imaging data. The whole channel number (All Channel) and the channel of interest (ChanOI) are required to be specified. Image files (.lsm or .tif) can be loaded by 'Load Image' button. The desired frame specified by a frame number ('Frame No' button) can be viewed by 'View' button.

4 Analysis of radial profiles

The position of the bleaching centre (X: column and Y: row) is required to locate a position for analysis. Radius controls the size of the analysed area. The pixel size is used to scale the x axis for the radial profile plot. Then, the radial profile can be analysed by clicking on the 'Radial profile Plot' button. Smooth range is used to

smooth the extracted raw radial profile, while the raw profile can also be presented by clicking on the 'Raw Profile' button.

5 Saved files

After the analysis, the FRAP curve data will be automatically saved to the anaFRAP folder, which is in the same folder as the input data. The four radial profiles of before bleaching, immediately after bleaching, after half of final recovery and final recovery will be saved to the same folder as the analysed FRAP curve data. The other figures in separate windows can be saved manually by using the 'save as' function.

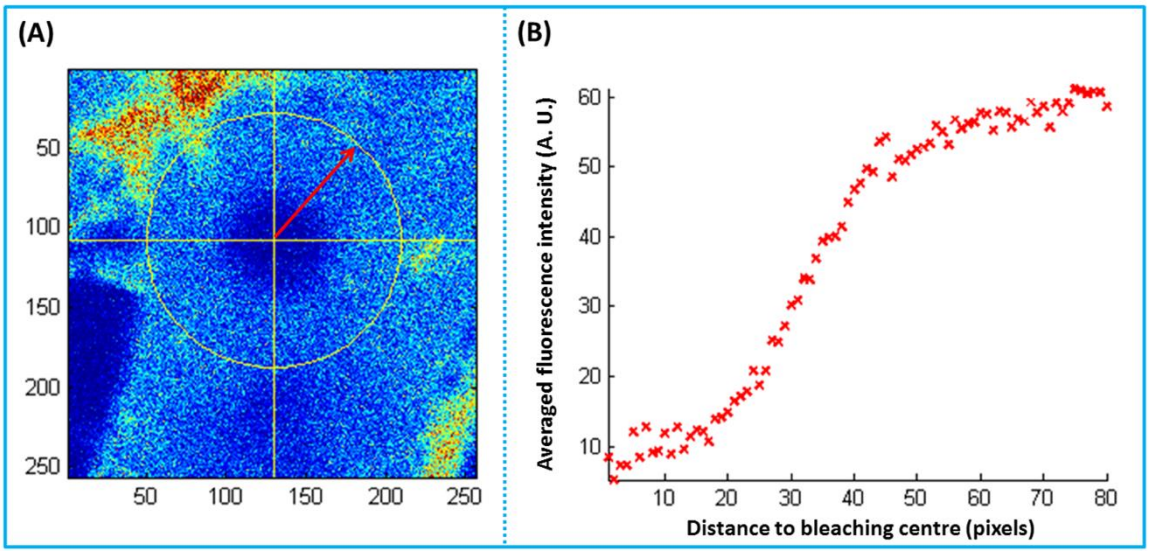


Figure 5.6 Radial profile analysis of an image. The radial profile of the image from the circle centre to the circle edge was extracted. (A): An image from the FRAP experiment; (B): The extracted radial profile of the fluorescence intensity.

5.5 Conclusion

In this chapter, the binding and the diffusion of five Halo-FGFs were studied by confocal microscopy. The methods for acquisition and processing of imaging data and their analysis with Matlab programs were developed, which will be useful to other researchers. There are five main findings revealed by this work. (1): HS is the main binding component in Rama 27 fibroblast pericellular matrix for FGF1, FGF2 and FGF6, whereas FGF10 can also bind substantially to CS/DS chains and FGF20 does not bind to glycosaminoglycans produced by these cells. (2): The binding sites for each FGF were heterogeneously distributed in the matrix, which is consistent with the previous study with gold nanoparticle-FGF2 (Duchesne et al., 2012) and suggests that this may be a more general property of pericellular matrix. (3): The number of binding sites in the matrix is different for each FGF. This suggests that the HS chains in the pericellular matrix have strong selectivity to these FGFs rather than binding to every paracrine FGF. (4): The FRAP measurements indicate that different FGFs move in the matrix with different speeds, which allows the formation of different paracrine protein gradients. (5): The competition experiments support the idea generated from the binding and the FRAP experiments, that the pericellular matrix possesses a degree of supramolecular organisation. The competition data also hint at this supramolecular structure being altered when an exogenous protein binds, exemplified by the strong competition by FGF2 for FGF1 (Section 5.3.3).

Chapter 6 General discussion

6.1 Using *E. coli* to make proteins

The development of recombinant protein expression has transformed our understanding of biomolecular systems and the ability to manipulate these for clinical purposes. In the past decades, many proteins have been produced in *E. coli*, because yields are higher and costs lower than in other hosts. This has allowed the investigation of protein structure, interactions with biological molecules and synthetic ligands, and the elucidation of their biological activities from the level of cells to that of tissues and the organism. Many of these protein studies were designed to explore the structures underpinning activities and interactions, which have then led to the screening and identification of small molecule inhibitors. For example, many small molecule kinase inhibitors (*e.g.*, imatinib and dasatinib, which interact with the kinase domain of Abelson murine leukemia viral oncogene homolog (ABL)) are used to treat cancers (Zhang et al., 2009). Moreover, the proteins themselves may have activities that are medically useful, exemplified by the use of FGFs in treating wounds and various ulcerative conditions (Nunes et al., 2015). So, it is very important to efficiently produce soluble and stable proteins with corresponding biological activities.

Central to this thesis was the production of a series of FGFs with well characterised HS binding properties to investigate the functional significance of their interactions with the HS co-receptor. In Chapter 3, HaloTag was successfully used as a solubilisation partner to express FGFs. HaloTag enabled a number of FGFs that were

expressed as insoluble proteins to be soluble; moreover, the yield of proteins was often increased. Some FGFs were still hardly expressed as soluble proteins (*e.g.*, Halo-FGF16) even with the HaloTag fusion. The yield of a protein is also an important issue, since a high yield of soluble protein makes purification simpler and also reduces the expense of protein production, which are important for both laboratory and large scale production of proteins. Thus, there is still considerable optimisation required for the expression of some of these FGFs. Another important aspect is the biological activities of the recombinant protein. The biological activities of produced FGFs were mainly tested by measuring the stimulation of MAPK phosphorylation. The cell signalling study indicates that the N-terminal HaloTag did not change the stimulation of phosphorylation of p44/42^{MAPK} by the FGFs, which suggests HaloTag has little effect on their biological activities. This is consistent with the fact that different FGFs possess N-terminal extensions that vary greatly in length (Fig. 1.6 and (Xu et al., 2012)) and with the observation that FGF2 with a gold nanoparticle of ~9 nm diameter conjugated to the N-terminus stimulated MAPK phosphorylation identically to native FGF2 (Duchesne et al., 2012). Since some proteins (Halo-FGF16, FGF17 and Halo-FGF17) did not cause obvious phosphorylation of p44/42^{MAPK}, a cell growth assay was used to test their biological activity. Cell growth is regulated by both cell division and cell survival, which is a more conventional and straightforward way to test the bioactivities of FGFs. The positive effect on cell growth indicates that these three proteins are also correctly folded, but in these cells they may mediate their effects by signalling pathways other than phosphorylation of p44/42^{MAPK}. These FGFs all bind to heparin, a property that depends on the correct folding of the protein, since the canonical heparin binding site is formed from amino acids that are distant in sequence. Indeed, a peptide

corresponding to the part of the canonical binding site of FGF2 that is contiguous in sequence has an affinity 1000-fold lower for heparin than FGF2 (Kinsella et al., 1998). Taken together with the biological activity of the FGFs measured here, these data indicate that the purified Halo-FGFs are correctly folded proteins. Whether HaloTag has any subtle effects on MAPK or other signalling pathways activated by FGFRs (Section 1.7) could be explored in the future by removal of the HaloTag with TEV protease.

For some Halo-FGFs, *e.g.*, Halo-FGF6, Halo-FGF8 and Halo-FGF22, HaloTag cannot be removed, since the released FGF always aggregates when TEV protease was used to cleave FGF from HaloTag. Thus, the work presented in Chapter 3 represents a partial solution to the challenge of producing high quality soluble recombinant FGFs.

A common strategy to produce soluble recombinant FGFs is to truncate the N- and sometimes the C-termini, leaving the core β trefoil core. However, these parts of the FGFs are likely to have important functions for two reasons. Firstly, there is considerable variability in these regions of FGFs, which is most pronounced between FGF subfamilies, whereas within a subfamily these regions are of more similar length (Fig. 1.6 and (Xu et al., 2012)), suggesting they have undergone the same natural selection processes that led to the diversification of the FGF family and its functions. Secondly, the fact that they are largely unstructured itself also supports this contention (Beenken et al., 2012, Moy et al., 1996, Olsen et al., 2006). The unstructured regions of proteins are usually involved in molecular interactions, and are involved in the molecular regulation of signalling processes (Wright and Dyson, 2015); in the case of the FGFs, there is evidence that this includes binding to the

FGFR and to HS *via* HBS3 for that part of the N-terminal unstructured region closest to strand β1 (Olsen et al., 2006, Beenken et al., 2012, Uniewicz et al., 2010, Xu et al., 2012). Thus, it will be important to develop means to produce full length soluble FGFs, including the more difficult cases. In some instances, although yields are far lower, mammalian expression systems may have to be employed.

Since some FGF2 medicines have already been used to accelerate wound healing, the applications of FGF medicines are constantly being extended. As increasing numbers of studies are focused on the biological functions of different FGFs from embryonic development to tissue repair and some diseases, successful production of FGFs will be of great importance, which may lead us to new FGF medicines in the future.

6.2 Binding selectivity of FGFs to heparin/HS

The analysis of the HS binding specificity of five FGFs from four subfamilies suggested that this had been subjected to the same natural selection pressure that drove the diversification of the 18 paracrine and endocrine FGFs, the specificity of their functions, and their binding to FGFR isoforms (Xu et al., 2012, Xu et al., 2013). The previous work had suggested that the binding specificity of FGFs for their sugar ligands (Table 6.1, red coloured FGFs), including heparin, heparin derivatives and short heparin oligosaccharides arose as part of the natural selection process that led to the expansion and diversification of the FGF family (Uniewicz et al., 2010, Xu et al., 2012). However, only two members from one subfamily (FGF1 and FGF2) were part of the analysis.

In the current work, this analysis was extended to include further members of the FGF7 and FGF9 subfamilies, FGF10 and FGF20 (Section 4.3). FGF9 and FGF20 are from same subfamily (FGF9 subfamily) and were found to possess very similar binding selectivity for structures in the library of polysaccharides (Table 6.1, green coloured FGFs). Both of these proteins need a longer heparin oligosaccharide (dp4 for FGF9 and dp12 for FGF20) for binding, which is one aspect that distinguishes them from FGFs in other subfamilies. FGF7 and FGF10 (from the FGF7 subfamily) bind to very similar heparin structures (Table 6.1, green coloured FGFs). They can also bind to DS and to CS, which is a unique property that distinguishes them from the other subfamilies; why this is the case is not known. Thus, the idea that the structures recognised by FGFs in the polysaccharide are more closely related for FGFs from the same subfamily, than for those from different subfamilies, a conclusion of the initial analysis (Xu et al., 2012) is supported by the present work.

Table 6.1 Binding selectivity of FGFs to heparin, heparin derivatives and heparin oligosaccharides (Uniewicz et al., 2010, Xu et al., 2012, Li et al., 2015). Red coloured FGFs are from previous publications; green FGFs are from Chapter 4 and blue FGFs were analysed by Yong Li (University of Liverpool) (Li, 2015).

Subfamily	Name	Heparin effect	Sulfation preference	dp length
FGF1	FGF1	> 22°C	NS ≈ 2S ≈ 6S	dp2 ~ dp10
	FGF2	> 22°C	NS, 2S > 6S	dp4 ~ dp10
FGF4	FGF4	~ 12°C	NS ≈ 2S > 6S	dp4 ~ dp10
	FGF6	> 10°C	2S > 6S, NS	dp6 ~ dp10
FGF7	FGF3	~ 15°C	NS ≈ 6S > 2S	dp4 ~ dp10
	FGF7	~ 7°C	NS ≈ 2S ≈ 6S	dp4 ~ dp8
	FGF10	~ 15°C	NS ≈ 2S ≈ 6S	dp4 ~ dp8
FGF8	FGF17	~ 15°C	NS ≈ 2S ≈ 6S	dp4 ~ dp10

	FGF18	> 15°C	NS ≈ 6S > 2S	dp6 ~ dp10
FGF9	FGF9	~ 20°C	NS ≈ 6S > 2S	dp4 ~ dp10
	FGF20	~ 12°C	6S ≈ 2S > NS	dp10 ~ dp12

Together with work performed in a parallel thesis (Li, 2015), there is now an almost fully comprehensive dataset on the interactions of FGFs with GAGs (Table 6.1). The hypothesis that the polysaccharide structures recognised by FGFs are more similar for FGFs in the same subfamily than for FGFs in different subfamilies is supported by these data (Table 6.1). This strongly suggests that the diversification of the FGF family and its functions in the course of evolution involved the diversification of the binding specificities of the FGFs for GAG co-receptor and FGFRs (Table 3.2 and (Zhang et al., 2006b, Ornitz et al., 1996)).

The binding experiments with Halo-FGF1 and Halo-FGF2 (Chapter 5) indicate that the differences in binding specificity between members of the same subfamily can have a major impact on their interaction with HS in pericellular matrix. Thus, not only are there more binding sites in pericellular matrix HS of Rama 27 cells for Halo-FGF1, which is consistent with its greater promiscuity, but it also moves more slowly than FGF2, which will affect its range and potential gradients. The similarities in binding and diffusion of Halo-FGF2 and Halo-FGF6 may also reflect their differences in binding specificities: FGF2 preferentially binds structures containing N- and 2-O- sulfate groups from dp4; FGF6 prefers structures containing 2-O- with either N- or 6-O- sulfate groups, but these need to be dp6 or longer. Thus, though FGF6 is more promiscuous in terms of sulfation pattern, it requires longer structures, which are rarer. In the case of Rama 27 cell pericellular matrix, these differences may compensate for each other, within the limits of detection and spatial resolution of the fluorescence measurements. The binding preferences for FGF10

and FGF20 are clearly more pronounced and this is too reflected in their binding and diffusion in Rama 27 pericellular matrix. Thus, at least for the FGFs investigated in Chapter 5, there is evidence that the divergence of the specificities of these FGFs for binding to GAG co-receptor may have significance at the level of controlling FGF movement between cells and this may have been one driver in the expansion of the FGF family.

Taken collectively the available data suggest that the interactions of FGFs with HS are both specific and selective, but there is no simple one-to-one correspondence between a particular FGF and a single sequence of saccharides. This is not surprising. The interactions of FGFs with isoforms of the FGFR is also not one-to-one, with the exception of that of FGF7 with FGFR2b (Zhang et al., 2006b). In more general terms, interactions of proteins with other polymers (DNA, RNA) generally demonstrates a graded specificity and selectivity: there is a consensus structure recognised by the protein, but more often than not the *in vivo* regulatory element recognised by the protein is some way off this consensus. This allows other interactions to participate in regulation and such "nonoptimal" (from an *in vitro* perspective) binding structures lie at the heart of the molecular regulation of many complex biological systems.

For HS there is a misconception, that proteins recognise a sequence of saccharides. This in part stems from the unusual pentasaccharide sequence that binds antithrombin III and so regulates coagulation. The polysaccharide is not linear in solution (no polymer is), but possesses a range of conformations. Substitution of a single sulfate group alters the conformation of the polysaccharide, and it is the three dimensional disposition of specific groups (sulfate, carboxyl and hydroxyl) that

enables protein binding. Clear evidence is provided by the identification of antithrombin III binding structures that are distinct from the pentasaccharide (Petitou et al., 1999b, Petitou et al., 1999a), and by the generation of FGF binding structures, able to support FGFR activation, from completely unrelated (from a saccharide backbone perspective) sulfated plant polysaccharides (Rudd et al., 2010). Thus, the key finding is that the molecular specificity of the interactions of FGFs with HS has followed the diversification of this growth factor family, which in turn reflects natural selection pressures that led to the formation of more complex animal body plans. This leads to the conclusion that the interactions of FGFs with HS are specific, selective and functionally important.

6.3 Binding and movement of FGFs in pericellular matrix

Given the above, it is clear that *in vitro* experiments that aim to elucidate the specificity and selectivity of FGFs for HS can only achieve this in part. The polysaccharide *in vivo* is far more complex than the purified structures used *in vitro*. For example, HS *in vivo* will have bound endogenous heparin binding proteins and coordinated cations, both of which may change the conformation of the HS chain. In addition, there are data that demonstrate clearly that HS in different tissue compartments has different protein binding properties (Rudland et al., 1993, Allen et al., 2001, Thompson et al., 2011). Though most of this work is at cellular resolution in tissues, at least for FGF2 it is also true at molecular length scales, since its binding sites are clustered in Rama 27 pericellular matrix (Duchesne et al., 2012). The binding and FRAP experiments (Chapter 5), therefore, probe far more than the number of saccharide binding sequences in HS and the diffusion of FGFs in pericellular matrix. If the FGF is considered as a (complex) reporter of HS structure,

it informs on how the binding structures on HS for that FGF are displayed, which in turn provides an indirect insight into the structure of the pericellular matrix. By using fixed Rama 27 fibroblasts, the selectivity of the pericellular matrix to different HS binding FGFs was studied with Halo-FGF1, Halo-FGF2, Halo-FGF6, Halo-FGF10 and Halo-FGF20 in the absence of confounding effects of cellular biochemistry.

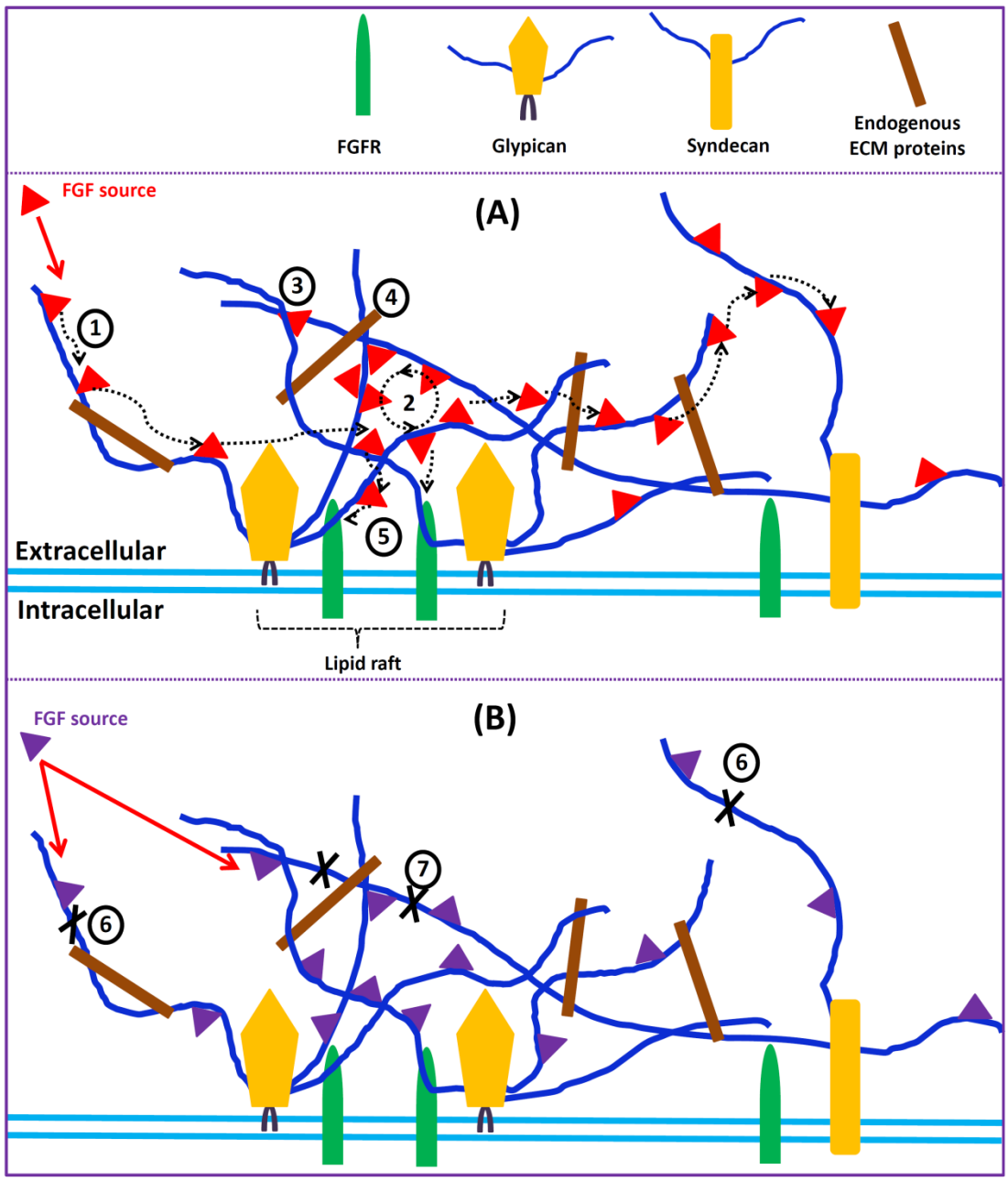


Figure 6.1 Model of binding and movement of FGFs in the pericellular matrix. (A): Model for mobile FGFs (*e.g.*, FGF1, FGF2 and FGF6). **(B):** Model for

immobile FGFs (*e.g.*, FGF10). The cell membrane contains two types of constitutive proteoglycans: glypican and syndecan. The pericellular matrix contains HS chains, plasma membrane proteins (*e.g.*, FGFR) endogenous matrix proteins (*e.g.*, collagens) and HS binding signalling proteins (*e.g.*, FGFs). The organisation of binding sites for FGFs is regulated by three key factors: organisation of membrane proteins (e.g., lipid rafts, links of plasma membrane proteins to cytoskeleton), the sulfation structure of HS (blue chains) and cross-linking of HS chains by interaction with endogenous ECM proteins. ①, diffusive movement of FGFs; ②, confined movement of FGFs and/or clustered binding; ③, cross-linking of HS chains by FGFs; ④, cross-linking of HS chains by endogenous ECM proteins (*e.g.*, collagen); ⑤, transport of FGF to FGFR; ⑥, release of FGFs by digestion of HS chains by heparanase; ⑦, digestion of HS chains without release of FGFs, because the HS chain is linked to a endogenous ECM protein.

The distribution of the Halo-FGFs in Rama 27 fibroblast pericellular matrix was heterogeneous, which indicates that their binding sites are not evenly distributed. This would arise from HS chains with different binding sites being spatially segregated, and to some of these binding sites being occupied by endogenous proteins, e.g., (Fig 6.1A ① to ④). CS and DS were also found to be involved in FGF binding and in organisation of pericellular matrix in two different ways. CS and DS had a direct role for FGF10, whereas they had an indirect role, for FGF1 and FGF6, since digestion of CS/DS chains was found to increase their binding (Section 5.2).

The data, therefore, show that the pericellular matrix can selectively bind different FGFs. Previous data led to the same conclusion (Rudland et al., 1993, Allen et al., 2001, Thompson et al., 2011), but the resolution was much coarser. Here the conclusion applies to a single cell. This provides strong support for the idea that

pericellular matrix is a highly organised structure, which allows the cells to control the number and spatial distribution of FGF binding sites (Fig. 6.1).

The data in this thesis also demonstrate that the control of binding and movement of FGFs in the pericellular matrix is not simply down to the cell synthesising HS chains with a particular sequence of saccharides and then localising the proteoglycan core protein to a particular place. For example, the differential effect of digestion of CS/DS with chondroitinase ABC on binding of FGF2 (no change), FGF1 and FGF6 (both increase in binding) seems likely to be due to a reorganisation of protein-polysaccharide interactions, which in some, but not all, cases frees up binding sites previously occupied or inaccessible. Similarly, the unexpected more effective competition by FGF2 for Halo-FGF1 binding, compared to competition by FGF1 cannot be explained by simple mass action driven competition, but seems likely to need changes in the organisation of HS chains driven by FGF binding. Therefore, the structure of pericellular matrix of Rama 27 fibroblasts is regulated by cross-linking of GAGs chains, especially those of HS, by endogenous proteins and the exogenously added FGF (Fig. 6.1 A, ③ and ④).

Different FGFs were found to move in the pericellular matrix with different speeds. FGFs that bind weakly or not at all to a particular configuration of HS, as in the case of FGF20 here, would be expected to undergo free diffusion in that locality, while FGFs, that do bind the HS chains may form gradients of different shapes (*e.g.*, long-slim gradient or short-fat gradients, Section 1.2) and also be stored in the matrix. Only a proportion of the FGFs (52% for FGF1, 81% for FGF2, 82% for FGF6 and none for FGF10) diffuses fast on the timescale of the FRAP experiments, and the rest seems to be strongly trapped (Figs 6.1 A and B, ②, ⑥ and ⑦). Although there are

no direct data, it seems reasonable to suppose that the immobile fraction in FRAP experiments is related to the confined motion (Fig. 6.1A, ②) observed by Duchesne *et al*. (Duchesne et al., 2012). Thus, the different classes of movement identified in the latter study, which additionally include diffusive motion and directed/restricted motion may be common to many FGFs and indeed to many HS-binding signalling proteins. What may differ is the proportion of each protein that is confined ("immobile" in the present work, Figs 6.1 A and B, ②, ⑥ and ⑦) or moving (Fig. 6.1 A, ①) and the relative proportion of the types of movement a protein can undergo.

In summary, how a matrix may selectively bind and transport individual HS-binding proteins will depend on different factors (Figs 6.1 A and B). In the fixed cell system, the structure of the pericellular matrix is organised by three main factors. Firstly, the location of the proteoglycan core protein, its association with cytoskeletal elements through its cytoplasmic domain (in the case of syndecans) and membrane microdomains, *e.g.*, rafts (glypicans), will also be important (Fig. 6.1 A, lipid raft) (Pike, 2004, Lingwood and Simons, 2010, Simons and Gerl, 2010). Secondly, the sequence of saccharides of the HS chain will play an important role, but it is clearly not the sole regulator (Figs 6.1, A and B, different binding sites for different FGFs). Thirdly, the interactions of the HS chains with endogenous cellular HS-binding proteins (Fig. 6.1 A, ④), some of which will be membrane bound and others not, and further interactions of these proteins with other elements, such as CS, DS and HA, will influence the conformation of the HS chains, which in turn may alter their binding properties. The actions of heparanase on HS chains and the degree to which the latter are cross-linked by endogenous proteins will also contribute to the control of movement of FGFs (Fig. 6.1 B). Thus, the evidence presented in this thesis

indicates that pericellular matrix has a supramoleclar structure, which selectively and heterogeneously binds to proteins and regulates the movement of these proteins to their target cells with different speeds. The live cell would be more complicated, since many cell activities (the dynamics of cell membrane and membrane proteins and cell endocytosis) will add a further layer of regulation and complexity (Belenkaya et al., 2004, Yu et al., 2009).

6.4 Suggestions on future work

Future work falls into two categories, that which is reasonably straightforward and can be explored in the immediate future and that which is more speculative. So far the interactions of only a limited number of FGFs with Rama 27 pericellular matrix have been explored. Thus, a HaloTag could be added to the N-terminus of FGFs that are currently untagged FGFs (FGF4, FGF5, FGF9 and FGF18). This may increase the soluble expression of these proteins and enable imaging and some biosensor studies that would expand on existing work on crosslinking of HS brushes (Migliorini et al., 2015) and mobility of the FGF in these models, very simplified extracellular matrices. For Halo-FGFs with low yield after heparin-affinity chromatography, re-engineering will be required, likely involving some degree of trimming of the N- and C-termini. It may also be necessary in some cases to produce the protein in mammalian systems or to engage in refolding of insoluble protein.

Since FGFs have different preferences for HS structure, an extended series of Halo-FGFs would be a very useful tool for the analysis of HS chains in matrix. In addition, it would be important to extend the present work in several directions. So far only the pericellular matrix of one cell line has been investigated. Working with an epithelial

cell line would likely be very informative, since the pericellular matrix will be different. Moreover, the height of the cells is sufficient to enable the lateral and basal pericellular matrices to be imaged independently of the apical one. More challenging would be to work in a 3-dimensional culture system. The microscopy technology now exists at the Liverpool Centre for Cell Imaging, which received one of the first Zeiss light sheet microscopes in the UK. However, how to make measurement is less clear. A FRAP type of experiment might be possible: bleaching a series of adjacent lines (the light sheets) by increasing the power of the laser, but the time resolution is lower than in confocal microscopy, so information will be more limited. An interesting alternative would be to pump the Halo-FGF into the bottom of the sample and thus provide the sphere of cells with a "source" of FGF and to measure the diffusion of the FGF into and through the cell spheroid. Again this is not trivial, since there may be 'leakage' of the FGF through the agarose gel in which the spheroid is embedded and this may confound measurements. Such experiments could also take on a pulse-chase format, using for example, first a red FGF then a green one.

The above analyses are limited by optical resolution and to understand the large scale organisation of matrix, molecular scale resolution would be an essential addition. Since the movement of FGFs is generally too fast for superresolution microscopy, the simplest solution would be to use electron microscopy. For this, a suitable chloromethyl terminated ligand for gold nanoparticles could be developed, which would allow the Halo-FGFs to be linked to gold nanoparticles. Alongside, antibody or the smaller nanobodies from camelids could be used to probe the location of various protein components, ranging from matrix proteins such as collagens, to the core proteins of proteoglycans.

Changing the structure of the HS on the cells would also be useful to understand how the sulfation modifications impact on HS chain and matrix structure. This would be best achieved by siRNA knockdown of HS biosynthetic enzymes, since it is rapid and so less likely to suffer from homeostatic mechanisms that compensate for the perturbation, as seen, for example, with the 2-O sulfotransferase knockout mouse (Bullock et al., 1998). However, siRNA is never completely effective and a more useful tool would be small molecule inhibitors of the sulfotransferases, though these do not exist yet.

Finally, there is now a considerable body of data on the binding and movement of FGFs in pericellular matrix. It should be possible, therefore, to generate some simple models to allow simulations. Some informal work with the Department of Mathematics has been undertaken and this is likely to yield in the future a reasonable working model. Such simulations would be an effective means to generate hypotheses, which could then be tested.

The Halo-FGFs would also be useful to probe the assembly of the ligand-receptor complex, since their colour can be altered by simply changing the fluorescent Halo-ligand they react with. By using a high concentration of competing oligosaccharide (as done in Duchesne *et al*., 2012, (Duchesne et al., 2012)), chlorate treatment (prevents all sulfation reaction) or siRNA for the EXTs (prevents HS biosynthesis), the exogenous oligosaccharide would be incorporated into the FGFR complex and would prevent binding of the FGF to any HS in the matrix. This would allow the assembly of the FGF onto the FGFR to be probed in the absence of the much greater signal (>99%) arising from the FGF binding HS. HaloTag may also be incorporated into the extracellular domain of the FGFR and this would enable ligand and FGFR to

be analysed independently, but simultaneously. Gene editing technology such as crisper would allow analysis of cells expressing normal levels of FGFR, which is likely to be important, since overexpression may lead to artificial associations of the FGFR. The use of gold nanoparticles with Halo ligands would then allow electron microscopy analysis, which would provide molecular scale validation of the optical measurements for FGF-FGFR assembly.

Contributions to following Papers and manuscripts

Sun, C., Li, Y., Taylor, S. E., Mao, X., Wilkinson, M. C., Fernig, D. G. (2015). HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors. *PeerJ*. **3**, e1060.

Contribution described in Chapter 3

Migliorini, E., Thakar, D., Kuhnle, J., Sadir, R., Dyer, D. P., Li, Y., **Sun, C.**, Volkman, B. F., Handel, T. M., Coche-Guerente, L., *et al.* (2015). Cytokines and growth factors cross-link heparan sulfate. *Open Biol.* **5**, 150046.

Provided FGF proteins for the experiments

Nunes, Q. M., Li, Y., **Sun, C.**, Kinnunen, T. K., Fernig, D. G. (2015). Fibroblast growth factors as tissue repair and regeneration therapeutics. *PeerJ PrePrints*. **3**, e1712.

Co-wrote the review and translated Chinese language papers to enable evaluation of clinical importnace by Quentin Nunes

Sun, C., Marcello, M., Li, Y., Mason, D., Levy, R., Fernig, D. G. (2015). Selectivity in glycosaminoglycan-binding dictates the distribution and diffusion of fibroblast growth factors in the pericellular matrix. (submitted to *Open Biology*)

Contribution described in Chapter 5

Sun, C., Li, Y., Yates, E. A., Fernig, D. G. (2015). SimpleDSFviewer: a tool to analyse and view differential scanning fluorimetry data for characterising protein thermal stability and interactions. PeerJ PrePrints 3:e1937.

Contribution described in Chapter 4

Li, Y., **Sun, C.**, Yates, E. A., Wilkinson, M. C., Fernig, D. G. (2015). Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification. (Submitted to Open Biology)

Provided Halo-FGF6, FGF10 and FGF20 proteins for some experiments and collected DSF data for FGF10 and FGF20; the latter data are described in Chapter 4; edited paper.

References

AKIYAMA, T. & GIBSON, M. C. 2015. Morphogen transport: theoretical and experimental controversies. *Wiley Interdisciplinary Reviews-Developmental Biology,* 4**,** 99-112.

ALLEN, B. L., FILLA, M. S. & RAPRAEGER, A. C. 2001. Role of heparan sulfate as a tissue-specific regulator of FGF-4 and FGF receptor recognition. *Journal of Cell Biology,* 155**,** 845-857.

ASADA, M., SHINOMIYA, M., SUZUKI, M., HONDA, E., SUGIMOTO, R., IKEKITA, M. & IMAMURA, T. 2009. Glycosaminoglycan affinity of the complete fibroblast growth factor family. *Biochim Biophys Acta,* 1790**,** 40-48.

BATTAGLIA, C., MAYER, U., AUMAILLEY, M. & TIMPL, R. 1992. Basement-membrane heparan sulfate proteoglycan binds to laminin by its heparan sulfate chains and to nidogen by sites in the protein core. *European Journal of Biochemistry,* 208**,** 359-366.

BEENKEN, A., ELISEENKOVA, A. V., IBRAHIMI, O. A., OLSEN, S. K. & MOHAMMADI, M. 2012. Plasticity in Interactions of Fibroblast Growth Factor 1 (FGF1) N Terminus with FGF Receptors Underlies Promiscuity of FGF1. *Journal of Biological Chemistry,* 287**,** 3067-3078.

BEENKEN, A. & MOHAMMADI, M. 2009. The FGF family: biology, pathophysiology and therapy. *Nature Reviews Drug Discovery,* 8**,** 235-253.

BELENKAYA, T. Y., HAN, C., YAN, D., OPOKA, R. J., KHODOUN, M., LIU, H. Z. & LIN, X. H. 2004. Drosophila Dpp morphogen movement is independent of dynamin-mediated endocytosis but regulated by the glypican members of heparan sulfate proteoglycans. *Cell,* 119**,** 231-244.

BERNFIELD, M., GOTTE, M., PARK, P. W., REIZES, O., FITZGERALD, M. L., LINCECUM, J. & ZAKO, M. 1999. Functions of cell surface heparan sulfate proteoglycans. *Annual Review of Biochemistry,* 68**,** 729-777.

BOKEL, C. & BRAND, M. 2013. Generation and interpretation of FGF morphogen gradients in vertebrates. *Current Opinion in Genetics & Development,* 23**,** 415-422.

BONNANS, C., CHOU, J. & WERB, Z. 2014. Remodelling the extracellular matrix in development and disease. *Nature Reviews Molecular Cell Biology,* 15**,** 786-801.

BOUKAMP, P., PETRUSSEVSKA, R. T., BREITKREUTZ, D., HORNUNG, J., MARKHAM, A. & FUSENIG, N. E. 1988. Normal keratinization in a

spontaneously immortalized aneuploid human keratinocyte cell line. *Journal of Cell Biology,* 106**,** 761-771.

BULLOCK, S. L., FLETCHER, J. M., BEDDINGTON, R. S. P. & WILSON, V. A. 1998. Renal agenesis in mice homozygous for a gene trap mutation in the gene encoding heparan sulfate 2-sulfotransferase. *Genes & Development,* 12**,** 1894-1906.

CAMPBELL, R. 2010. *shadedErrorBar* [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/26311-shadederrorbar [Accessed 04 March 2015].

CASCI, T., VINOS, J. & FREEMAN, M. 1999. Sprouty, an intracellular inhibitor of Ras signaling. *Cell,* 96**,** 655-665.

CASTELLOT, J. J., WONG, K., HERMAN, B., HOOVER, R. L., ALBERTINI, D. F., WRIGHT, T. C., CALEB, B. L. & KARNOVSKY, M. J. 1985. Binding and Internalization of Heparin by Vascular Smooth-Muscle Cells. *Journal of Cellular Physiology,* 124**,** 13-20.

CHANG, Z., MEYER, K., RAPRAEGER, A. C. & FRIEDL, A. 2000. Differential ability of heparan sulfate proteoglycans to assemble the fibroblast growth factor receptor complex in situ. *Faseb Journal,* 14**,** 137-144.

CHATTRAPIBAN, N. 2007. *Radial Scan* [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/18102-radial-scan [Accessed 03 June 2014].

CHEN, C. L., YANG, J. X., JAMES, I. O. A., ZHANG, H. Y. & BESNER, G. E. 2014. Heparin-binding epidermal growth factor-like growth factor restores Wnt/beta-catenin signaling in intestinal stem cells exposed to ischemia/reperfusion injury. *Surgery,* 155**,** 1069-1080.

CHU, C. L., BUCZEK-THOMAS, J. A. & NUGENT, M. A. 2004. Heparan sulphate proteoglycans modulate fibroblast growth factor-2 binding through a lipid raft-mediated mechanism (vol 379, pg 331, 2004). *Biochemical Journal,* 380**,** 951-951.

CONNELL, B. J. & LORTAT-JACOB, H. 2013. Human immunodeficiency virus and heparan sulfate: from attachment to entry inhibition. *Frontiers in Immunology,* 4**,** doi: 10.3389/fimmu.2013.00385.

COUCHMAN, J. R. 2003. Syndecans: Proteoglycan regulators of cell-surface microdomains? *Nature Reviews Molecular Cell Biology,* 4**,** 926-937.

DAVIES, J. A. 2001. Extracellular matrix. *eLS.***,** 7.

DELEHEDDE, M., LYON, M., GALLAGHER, J. T., RUDLAND, P. S. & FERNIG, D. G. 2002a. Fibroblast growth factor-2 binds to small heparin-derived oligosaccharides and stimulates a sustained phosphorylation of p42/44 mitogen-activated protein kinase and proliferation of rat mammary fibroblasts. *Biochemical Journal,* 366**,** 235-244.

DELEHEDDE, M., LYON, M., VIDYASAGAR, R., MCDONNELL, T. J. & FERNIG, D. G. 2002b. Hepatocyte growth factor/scatter factor binds to small heparin-derived oligosaccharides and stimulates the proliferation of human HaCaT keratinocytes. *Journal of Biological Chemistry,* 277**,** 12456-12462.

DELEHEDDE, M., SEVE, M., SERGEANT, N., WARTELLE, I., LYON, M., RUDLAND, P. S. & FERNIG, D. G. 2000. Fibroblast growth factor-2 stimulation of p42/44(MAPK) phosphorylation and I kappa B degradation is regulated by heparan sulfate/heparin in rat mammary fibroblasts. *Journal of Biological Chemistry,* 275**,** 33905-33910.

DOREY, K. & AMAYA, E. 2010. FGF signalling: diverse roles during early vertebrate embryogenesis. *Development,* 137**,** 3731-3742.

DOWD, C. J., COONEY, C. L. & NUGENT, M. A. 1999. Heparan sulfate mediates bFGF transport through basement membrane by diffusion with rapid reversible binding. *Journal of Biological Chemistry,* 274**,** 5236-5244.

DREYFUSS, J. L., REGATIERI, C. V., JARROUGE, T. R., CAVALHEIRO, R. P., SAMPAIO, L. O. & NADER, H. B. 2009. Heparan sulfate proteoglycans: structure, protein interactions and cell signaling. *Anais Da Academia Brasileira De Ciencias,* 81**,** 409-429.

DUCHESNE, L., GENTILI, D., COMES-FRANCHINI, M. & FERNIG, D. G. 2008. Robust ligand shells for biological applications of gold nanoparticles. *Langmuir,* 24**,** 13572-13580.

DUCHESNE, L., OCTEAU, V., BEARON, R. N., BECKETT, A., PRIOR, I. A., LOUNIS, B. & FERNIG, D. G. 2012. Transport of fibroblast growth factor 2 in the pericellular matrix is controlled by the spatial distribution of its binding sites in heparan sulfate. *Plos Biology,* 10**,** e1001976.

EBLAGHIE, M. C., LUNN, J. S., DICKINSON, R. J., MUNSTERBERG, A. E., SANZ-EZQUERRO, J. J., FARRELL, E. R., MATHERS, J., KEYSE, S. M., STOREY, K. & TICKLE, C. 2003. Negative feedback regulation of FGF signaling levels by Pyst1/MKP3 in chick embryos. *Current Biology,* 13**,** 1009-1018.

FAHAM, S., HILEMAN, R. E., FROMM, J. R., LINHARDT, R. J. & REES, D. C. 1996. Heparin structure and interactions with basic fibroblast growth factor. *Science,* 271**,** 1116-1120.

FAN, D., CREEMERS, E. E. & KASSIRI, Z. 2014. Matrix as an interstitial transport system. *Circulation Research,* 114**,** 889-902.

FAN, H. H., VITHARANA, S. N., CHEN, T., O'KEEFE, D. & MIDDAUGH, C. R. 2007. Effects of pH and polyanions on the thermal stability of fibroblast growth factor 20. *Molecular Pharmaceutics,* 4**,** 232-240.

FERNIG, D. G., SMITH, J. A. & RUDLAND, P. S. 1990. Appearance of basic fibroblast growth factor receptors upon differentiation of rat mammary

epithelial to myoepithelial-like cells in culture. *Journal of Cellular Physiology,* 142**,** 108-116.

FORD-PERRISS, M., GUIMOND, S. E., GREFERATH, U., KITA, M., GROBE, K., HABUCHI, H., KIMATA, K., ESKO, J. D., MURPHY, M. & TURNBULL, J. E. 2002. Variant heparan sulfates synthesized in developing mouse brain differentially regulate FGF signaling. *Glycobiology,* 12**,** 721-727.

FRANTZ, C., STEWART, K. M. & WEAVER, V. M. 2010. The extracellular matrix at a glance. *Journal of Cell Science,* 123**,** 4195-4200.

FRIEDL, A., CHANG, Z., TIERNEY, A. & RAPRAEGER, A. C. 1997. Differential binding of fibroblast growth factor-2 and -7 to basement membrane heparan sulfate - Comparison of normal and abnormal human tissues. *American Journal of Pathology,* 150**,** 1443-1455.

GOEL, S. & GOWDA, D. C. 2011. How specific is Plasmodium falciparum adherence to chondroitin 4-sulfate? *Trends in Parasitology,* 27**,** 375-381.

GOETZ, R. & MOHAMMADI, M. 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology,* 14**,** 166-180.

GONZALEZ, A. M., HILL, D. J., LOGAN, A., MAHER, P. A. & BAIRD, A. 1996. Distribution of fibroblast growth factor (FGF)-2 and FGF receptor-1 messenger RNA expression and protein presence in the mid-trimester human fetus. *Pediatric Research,* 39**,** 375-385.

GUGLIERI, S., HRICOVINI, M., RAMAN, R., POLITO, L., TORRI, G., CASU, B., SASISEKHARAN, R. & GUERRINI, M. 2008. Minimum FGF2 binding structural requirements of heparin and heparan sulfate oligosaccharides as determined by NMR spectroscopy. *Biochemistry,* 47**,** 13862-13869.

GUIMOND, S. E., PUVIRAJESINGHE, T. M., SKIDMORE, M. A., KALUS, I., DIERKS, T., YATES, E. A. & TURNBULL, J. E. 2009. Rapid purification and high sensitivity analysis of heparan sulfate from cells and tissues - TOWARD GLYCOMICS PROFILING. *Journal of Biological Chemistry,* 284**,** 25714-25722.

GURDON, J. B. & BOURILLOT, P. Y. 2001. Morphogen gradient interpretation. *Nature,* 413**,** 797-803.

HAAS, R. & CULP, L. A. 1984. Binding of fibronectin to gelatin and heparin - effect of surface denaturation and detergents. *Febs Letters,* 174**,** 279-283.

HARADA, H., KETTUNEN, P., JUNG, H. S., MUSTONEN, T., WANG, Y. A. & THESLEFF, I. 1999. Localization of putative stem cells in dental epithelium and their association with notch and FGF signaling. *Journal of Cell Biology,* 147**,** 105-120.

HARADA, M., MURAKAMI, H., OKAWA, A., OKIMOTO, N., HIRAOKA, S., NAKAHARA, T., AKASAKA, R., SHIRAISHI, Y., FUTATSUGI, N., MIZUTANI-KOSEKI, Y., KUROIWA, A., SHIROUZU, M., YOKOYAMA, S., TAIJI, M., ISEKI, S., ORNITZ, D. M. & KOSEKI, H. 2009. FGF9 monomer-dimer equilibrium regulates extracellular matrix affinity and tissue diffusion. *Nature Genetics,* 41**,** 289-298.

HARISI, R. & JENEY, A. 2015. Extracellular matrix as target for antitumor therapy. *Oncotargets and Therapy,* 8**,** 1387-1398.

HECHT, H. J., ADAR, R., HOFMANN, B., BOGIN, O., WEICH, H. & YAYON, A. 2001. Structure of fibroblast growth factor 9 shows a symmetric dimer with unique receptor- and heparin-binding interfaces. *Acta Crystallographica Section D-Biological Crystallography,* 57**,** 378-384.

HEINEGARD, D. & SAXNE, T. 2011. The role of the cartilage matrix in osteoarthritis. *Nature Reviews Rheumatology,* 7**,** 50-56.

IGARASHI, M., FINCH, P. W. & AARONSON, S. A. 1998. Characterization of recombinant human fibroblast growth factor (FGF)-10 reveals functional similarities with keratinocyte growth factor (FGF-7). *Journal of Biological Chemistry,* 273**,** 13230-13235.

ISRAELI, D., BENCHAOUIR, R., ZIAEI, S., RAMEAU, P., GRUSZCZYNSKI, C., PELTEKIAN, E., DANOS, O. & GARCIA, L. 2004. FGF6 mediated expansion of a resident subset of cells with SP phenotype in the C2C12 myogenic line. *Journal of Cellular Physiology,* 201**,** 409-419.

ITOH, N. 2007. The FGF families in humans, mice, and zebrafish: Their evolutional processes and roles in development, metabolism, and disease. *Biological & Pharmaceutical Bulletin,* 30**,** 1819-1825.

ITOH, N. & ORNITZ, D. M. 2004. Evolution of the Fgf and Fgfr gene families. *Trends in Genetics,* 20**,** 563-569.

ITOH, N. & ORNITZ, D. M. 2011. Fibroblast growth factors: from molecular evolution to roles in development, metabolism and disease. *Journal of Biochemistry,* 149**,** 121-130.

IZVOLSKY, K. I., SHOYKHET, D., YANG, Y., YU, Q., NUGENT, M. A. & CARDOSO, W. V. 2003. Heparan sulfate-FGF10 interactions during lung morphogenesis. *Developmental Biology,* 258**,** 185-200.

JEMTH, P., KREUGER, J., KUSCHE-GULLBERG, M., STURIALE, L., GIMENEZ-GALLEGO, G. & LINDAHL, U. 2002. Biosynthetic oligosaccharide libraries for identification of protein-binding heparan sulfate motifs - Exploring the structural diversity by screening for fibroblast growth factor (FGF) 1 and FGF2 binding. *Journal of Biological Chemistry,* 277**,** 30567-30573.

JOHNSON, D. E., LU, J., CHEN, H., WERNER, S. & WILLIAMS, L. T. 1991. The human fibroblast growth factor receptor genes - a common structural

arrangement underlies the mechanisms for generating receptor forms that differ in their 3rd immunoglobulin domain. *Molecular and Cellular Biology,* 11**,** 4627-4634.

KALININA, J., BYRON, S. A., MAKARENKOVA, H. P., OLSEN, S. K., ELISEENKOVA, A. V., LAROCHELLE, W. J., DHANABAL, M., BLAIS, S., ORNITZ, D. M., DAY, L. A., NEUBERT, T. A., POLLOCK, P. M. & MOHAMMADI, M. 2009. Homodimerization controls the fibroblast growth factor 9 subfamily's receptor binding and heparan sulfate-dependent diffusion in the extracellular matrix. *Molecular and Cellular Biology,* 29**,** 4663-4678.

KAN, M., WANG, F., KAN, M., TO, B., GABRIEL, J. L. & MCKEEHAN, W. L. 1996. Divalent cations and heparin/heparan sulfate cooperate to control assembly and activity of the fibroblast growth factor receptor complex. *Journal of Biological Chemistry,* 271**,** 26143-26148.

KATO, M., WANG, H. M., KAINULAINEN, V., FITZGERALD, M. L., LEDBETTER, S., ORNITZ, D. M. & BERNFIELD, M. 1998. Physiological degradation converts the soluble syndecan-1 ectodomain from an inhibitor to a potent activator of FGF-2. *Nature Medicine,* 4**,** 691-697.

KE, Y. Q., FERNIG, D. G., SMITH, J. A., WILKINSON, M. C., ANANDAPPA, S. Y., RUDLAND, P. S. & BARRACLOUGH, R. 1990. High level production of human acidic fibroblast growth factor in Escherichia coli cells inhibition of DNA synthesis in Rat mammary fibroblasts at high concentrations of growth factor. *Biochemical and Biophysical Research Communications,* 171**,** 963-971.

KE, Y. Q., WILKINSON, M. C., FERNIG, D. G., SMITH, J. A., RUDLAND, P. S. & BARRACLOUGH, R. 1992. A rapid procedure for production of human basic fibroblast growth factor in Escherichia Coli cells. *Biochimica Et Biophysica Acta,* 1131**,** 307-310.

KICHEVA, A., PANTAZIS, P., BOLLENBACH, T., KALAIDZIDIS, Y., BITTIG, T., JULICHER, F. & GONZALEZ-GAITAN, M. 2007. Kinetics of morphogen gradient formation. *Science,* 315**,** 521-525.

KIM, I., MOON, S. O., YU, K. H., KIM, U. H. & KOH, G. Y. 2001. A novel fibroblast growth factor receptor-5 preferentially expressed in the pancreas. *Biochimica Et Biophysica Acta-Gene Structure and Expression,* 1518**,** 152-156.

KINSELLA, L., CHEN, H. L., SMITH, J. A., RUDLAND, P. S. & FERNIG, D. G. 1998. Interactions of putative heparin-binding domains of basic fibroblast growth factor and its receptor, FGFR-1, with heparin using synthetic peptides. *Glycoconjugate Journal,* 15**,** 419-422.

KIRKPATRICK, C. A. & SELLECK, S. B. 2007. Heparan sulfate proteoglycans at a glance. *Journal of Cell Science,* 120**,** 1829-1832.

KLEINSCHMIT, A., KOYAMA, T., DEJIMA, K., HAYASHI, Y., KAMIMURA, K. & NAKATO, H. 2010. Drosophila heparan sulfate 6-O endosulfatase

regulates Wingless morphogen gradient formation. *Developmental Biology,* 345**,** 204-214.

KREUGER, J., JEMTH, P., SANDERS-LINDBERG, E., ELIAHU, L., RON, D., BASILICO, C., SALMIVIRTA, M. & LINDAHL, U. 2005. Fibroblast growth factors share binding sites in heparan sulphate. *Biochemical Journal,* 389**,** 145-150.

KREUGER, J., SPILLMANN, D., LI, J. P. & LINDAHL, U. 2006. Interactions between heparan sulfate and proteins: the concept of specificity. *Journal of Cell Biology,* 174**,** 323-327.

LAMOUILLE, S., XU, J. & DERYNCK, R. 2014. Molecular mechanisms of epithelial-mesenchymal transition. *Nature Reviews Molecular Cell Biology,* 15**,** 178-196.

LEDIN, J., STAATZ, W., LI, J. P., GOTTE, M., SELLECK, S., KJELLEN, L. & SPILLMANN, D. 2004. Heparan sulfate structure in mice with genetically modified heparan sulfate production. *Journal of Biological Chemistry,* 279**,** 42732-42741.

LEWANDOSKI, M., SUN, X. & MARTIN, G. R. 2000. Fgf8 signalling from the AER is essential for normal limb development. *Nature Genetics,* 26**,** 460-463.

LI, Y. 2015. *Specificities of the interaction of fibroblast growth factor and heparan sulfate.* PhD, University of Liverpool.

LI, Y., SUN, C., YATES, E. A., WILKINSON, M. C. & FERNIG, D. G. 2015. Heparin binding preference and structures in the fibroblast growth factor family parallel their evolutionary diversification. *Open Biol.***,** Submitted.

LIN, X. H. 2004. Functions of heparan sulfate proteoglycans in cell signaling during development. *Development,* 131**,** 6009-6021.

LINGWOOD, D. & SIMONS, K. 2010. Lipid rafts as a membrane-organizing principle. *Science,* 327**,** 46-50.

LIU, A. M. & JOYNER, A. L. 2001. Early anterior/posterior patterning of the midbrain and cerebellum. *Annual Review of Neuroscience,* 24**,** 869-896.

LORTAT-JACOB, H., GROSDIDIER, A. & IMBERTY, A. 2002. Structural diversity of heparan sulfate binding domains in chemokines. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 1229-1234.

LOS, G. V., ENCELL, L. P., MCDOUGALL, M. G., HARTZELL, D. D., KARASSINA, N., ZIMPRICH, C., WOOD, M. G., LEARISH, R., OHANE, R. F., URH, M., SIMPSON, D., MENDEZ, J., ZIMMERMAN, K., OTTO, P., VIDUGIRIS, G., ZHU, J., DARZINS, A., KLAUBERT, D. H., BULLEIT, R. F. & WOOD, K. V. 2008. HatoTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology,* 3**,** 373-382.

LU, P. F., WEAVER, V. M. & WERB, Z. 2012. The extracellular matrix: A dynamic niche in cancer progression. *Journal of Cell Biology,* 196**,** 395-406.

LUO, Y. D., YE, S., KAN, M. & MCKEEHAN, W. L. 2006. Structural specificity in a FGF7-affinity purified heparin octasaccharide required for formation of a complex with FGF7 and FGFR2IIIb. *Journal of Cellular Biochemistry,* 97**,** 1241-1258.

MACCARANA, M., CASU, B. & LINDAHL, U. 1993. Minimal sequence in heparin heparan-sulfate required for binding of basic fibroblast growth factor. *Journal of Biological Chemistry,* 268**,** 23898-23905.

MAITY, H., KARKARIA, C. & DAVAGNINO, J. 2009. Effects of pH and arginine on the solubility and stability of a therapeutic protein (fibroblast growth factor 20): relationship between solubility and stability. *Current Pharmaceutical Biotechnology,* 10**,** 609-625.

MAKARENKOVA, H. P., HOFFMAN, M. P., BEENKEN, A., ELISEENKOVA, A. V., MEECH, R., TSAU, C., PATEL, V. N., LANG, R. A. & MOHAMMADI, M. 2009. Differential interactions of FGFs with heparan sulfate control gradient formation and branching morphogenesis. *Science Signaling,* 2**,** ra55.

MATHWORKS. 2014. *Detecting a cell using image segmentation* [Online]. Available: http://uk.mathworks.com/help/images/examples/detecting-a-cell-using-image-segmentation.html [Accessed 21 January 2015].

MATHWORKS. 2015. *smooth (Smooth response data)* [Online]. Available: http://uk.mathworks.com/help/curvefit/smooth.html [Accessed 13 March 2015.

MATUS, D. Q., THOMSEN, G. H. & MARTINDALE, M. Q. 2007. FGF signaling in gastrulation and neural development in Nematostella vectensis, an anthozoan cnidarian. *Development Genes and Evolution,* 217**,** 137-148.

MCKEEHAN, W. L., WANG, F. & KAN, M. 1998. The heparan sulfate fibroblast growth factor family: Diversity of structure and function. *Progress in Nucleic Acid Research and Molecular Biology,* 59**,** 135-176.

MIGLIORINI, E., THAKAR, D., KUHNLE, J., SADIR, R., DYER, D. P., LI, Y., SUN, C., VOLKMAN, B. F., HANDEL, T. M., COCHE-GUERENTE, L., FERNIG, D. G., LORTAT-JACOB, H. & RICHTER, R. P. 2015. Cytokines and growth factors cross-link heparan sulfate. *Open Biol.,* 5**,** 150046.

MOHAMMADI, M., OLSEN, S. K. & IBRAHIMI, O. A. 2005. Structural basis for fibroblast growth factor receptor activation. *Cytokine & Growth Factor Reviews,* 16**,** 107-137.

MORI, S., KIUCHI, S., OUCHI, A., HASE, T. & MURASE, T. 2014. Characteristic expression of extracellular matrix in subcutaneous adipose tissue development and adipogenesis; comparison with visceral adipose tissue. *International Journal of Biological Sciences,* 10**,** 825-833.

MOUW, J. K., OU, G. Q. & WEAVER, V. M. 2014. Extracellular matrix assembly: a multiscale deconstruction. *Nature Reviews Molecular Cell Biology,* 15**,** 771-785.

MOY, F. J., SEDDON, A. P., BOHLEN, P. & POWERS, R. 1996. High-resolution solution structure of basic fibroblast growth factor determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry,* 35**,** 13552-13561.

MULLER, P., ROGERS, K. W., YU, S. Z. R., BRAND, M. & SCHIER, A. F. 2013. Morphogen transport. *Development,* 140**,** 1621-1638.

MUNDHENKE, C., MEYER, K., DREW, S. & FRIEDL, A. 2002. Heparan sulfate proteoglycans as regulators of fibroblast growth factor-2 receptor binding in breast carcinomas. *American Journal of Pathology,* 160**,** 185-194.

MURPHY, K. J., MERRY, C. L. R., LYON, M., THOMPSON, J. E., ROBERTS, I. S. & GALLAGHER, J. T. 2004. A new model for the domain structure of heparan sulfate based on the novel specificity of K5 lyase. *Journal of Biological Chemistry,* 279**,** 27239-27245.

NIESEN, F. H., BERGLUND, H. & VEDADI, M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols,* 2**,** 2212-2221.

NIEVES, D. J., LI, Y., FERNIG, D. G. & LEVY, R. 2015. Photothermal raster image correlation spectroscopy of gold nanoparticles in solution and on live cells. *R. Soc. open sci.,* 2**,** 140454.

NUNES, Q. M. 2015. *The role of heparin-binding proteins in normal pancreas and acute pancreatitis (PhD Thesis).* PhD, University of Liverpool.

NUNES, Q. M., LI, Y., SUN, C., KINNUNEN, T. K. & FERNIG, D. G. 2015. Fibroblast growth factors as tissue repair and regeneration therapeutics. *PeerJ PrePrints,* 3**,** e1712.

NUR-E-KAMAL, A., AHMED, I., KAMAL, J., BABU, A. N., SCHINDLER, M. & MEINERS, S. 2008. Covalently attached FGF-2 to three-dimensional polyamide nanofibrillar surfaces demonstrates enhanced biological stability and activity. *Molecular and Cellular Biochemistry,* 309**,** 157-166.

OLCZYK, P., MENCNER, L. & KOMOSINSKA-VASSEV, K. 2014. The role of the extracellular matrix components in cutaneous wound healing. *Biomed Research International,* 2014**,** Article ID 747584.

OLSEN, S. K., LI, J. Y. H., BROMLEIGH, C., ELISEENKOVA, A. V., IBRAHIMI, O. A., LAO, Z. M., ZHANG, F. M., LINHARDT, R. J., JOYNER, A. L. & MOHAMMADI, M. 2006. Structural basis by which alternative splicing modulates the organizer activity of FGF8 in the brain. *Genes & Development,* 20**,** 185-198.

ORI, A., FREE, P., COURTY, J., WILKINSON, M. C. & FERNIG, D. G. 2009. Identification of heparin-binding sites in proteins by selective labeling. *Molecular & Cellular Proteomics,* 8**,** 2256-2265.

ORI, A., WILKINSON, M. C. & FERNIG, D. G. 2008. The heparanome and regulation of cell function: structures, functions and challenges. *Frontiers in Bioscience,* 13**,** 4309-4338.

ORI, A., WILKINSON, M. C. & FERNIG, D. G. 2011. A systems biology approach for the investigation of the heparin/heparan sulfate interactome. *Journal of Biological Chemistry,* 286**,** 19892-19904.

ORNITZ, D. M. 2000. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays,* 22**,** 108-112.

ORNITZ, D. M. & ITOH, N. 2001. Fibroblast growth factors. *Genome Biology,* 2**,** REVIEWS3005.

ORNITZ, D. M. & MARIE, P. J. 2002. FGF signaling pathways in endochondral and intramembranous bone development and human genetic disease. *Genes & Development,* 16**,** 1446-1465.

ORNITZ, D. M., XU, J. S., COLVIN, J. S., MCEWEN, D. G., MACARTHUR, C. A., COULIER, F., GAO, G. X. & GOLDFARB, M. 1996. Receptor specificity of the fibroblast growth factor family. *Journal of Biological Chemistry,* 271**,** 15292-15297.

ORRURTREGER, A., BEDFORD, M. T., BURAKOVA, T., ARMAN, E., ZIMMER, Y., YAYON, A., GIVOL, D. & LONAI, P. 1993. Developmental localization of the splicing alternatives of fibroblast growth factor receptor-2 (FGFR2). *Developmental Biology,* 158**,** 475-486.

PANAKOVA, D., SPRONG, H., MAROIS, E., THIELE, C. & EATON, S. 2005. Lipoprotein particles are required for Hedgehog and Wingless signalling. *Nature,* 435**,** 58-65.

PANTOLIANO, M. W., PETRELLA, E. C., KWASNOSKI, J. D., LOBANOV, V. S., MYSLIK, J., GRAF, E., CARVER, T., ASEL, E., SPRINGER, B. A., LANE, P. & SALEMME, F. R. 2001. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *Journal of Biomolecular Screening,* 6**,** 429-440.

PATEL, V. N., KNOX, S. M., LIKAR, K. M., LATHROP, C. A., HOSSAIN, R., EFTEKHARI, S., WHITELOCK, J. M., ELKIN, M., VLODAVSKY, I. & HOFFMAN, M. P. 2007. Heparanase cleavage of perlecan heparan sulfate modulates FGF10 activity during ex vivo submandibular gland branching morphogenesis. *Development,* 134**,** 4177-4186.

PATEL, V. N., LIKAR, K. M., ZISMAN-ROZEN, S., COWHERD, S. N., LASSITER, K. S., SHER, I., YATES, E. A., TURNBULL, J. E., RON, D. & HOFFMAN, M. P. 2008. Specific heparan sulfate structures modulate

FGF10-mediated submandibular gland epithelial morphogenesis and differentiation. *Journal of Biological Chemistry,* 283**,** 9308-9317.

PELLEGRINI, L., BURKE, D. F., VON DELFT, F., MULLOY, B. & BLUNDELL, T. L. 2000. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature,* 407**,** 1029-1034.

PETITOU, M., DUCHAUSSOY, P., DRIGUEZ, P. A., HERAULT, J. P., LORMEAU, J. C. & HERBERT, J. M. 1999a. New synthetic heparin mimetics able to inhibit thrombin and factor Xa. *Bioorganic & Medicinal Chemistry Letters,* 9**,** 1155-1160.

PETITOU, M., HERAULT, L. P., BERNAT, A., DRIGUEZ, P. A., DUCHAUSSOY, P., LORMEAU, J. C. & HERBERT, J. M. 1999b. Synthesis of thrombin-inhibiting heparin mimetics without side effects. *Nature,* 398**,** 417-422.

PIKE, L. J. 2004. Lipid rafts: heterogeneity on the high seas. *Biochemical Journal,* 378**,** 281-292.

PLOTNIKOV, A. N., HUBBARD, S. R., SCHLESSINGER, J. & MOHAMMADI, M. 2000. Crystal structures of two FGF-FGFR complexes reveal the determinants of ligand-receptor specificity. *Cell,* 101**,** 413-424.

POWELL, A. K., FERNIG, D. G. & TURNBULL, J. E. 2002. Fibroblast growth factor receptors 1 and 2 interact differently with heparin/heparan sulfate - Implications for dynamic assembly of a ternary signaling complex. *Journal of Biological Chemistry,* 277**,** 28554-28563.

POWERS, C. J., MCLESKEY, S. W. & WELLSTEIN, A. 2000. Fibroblast growth factors, their receptors and signaling. *Endocrine-Related Cancer,* 7**,** 165-197.

QU, X. X., PAN, Y., CARBE, C., POWERS, A., GROBE, K. & ZHANG, X. 2012. Glycosaminoglycan-dependent restriction of FGF diffusion is necessary for lacrimal gland development. *Development,* 139**,** 2730-2739.

QUAGLINO, D., FORNIERI, C., NANNEY, L. B. & DAVIDSON, J. M. 1993. Extracellular matrix modifications in Rat tissues of different ages: Correlations between elastin and collagen type I mRNA expression and lysyl-oxidase activity. *Matrix,* 13**,** 481-490.

RADISKY, D. C. 2005. Epithelial-mesenchymal transition. *Journal of Cell Science,* 118**,** 4325-4326.

RAHMOUNE, H., CHEN, H. L., GALLAGHER, J. T., RUDLAND, P. S. & FERNIG, D. G. 1998. Interaction of heparan sulfate from mammary cells with acidic fibroblast growth factor (FGF) and basic FGF - Regulation of the activity of basic FGF by high and low affinity binding sites in heparan sulfate. *Journal of Biological Chemistry,* 273**,** 7303-7310.

RAPSOMANIKI, M. A., KOTSANTIS, P., SYMEONIDOU, I. E., GIAKOUMAKIS, N. N., TARAVIRAS, S. & LYGEROU, Z. 2012.

easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data. *Bioinformatics,* 28**,** 1800-1801.

RUDD, T. R., GUIMOND, S. E., SKIDMORE, M. A., DUCHESNE, L., GUERRINI, M., TORRI, G., COSENTINO, C., BROWN, A., CLARKE, D. T., TURNBULL, J. E., FERNIG, D. G. & YATES, E. A. 2007. Influence of substitution pattern and cation binding on conformation and activity in heparin derivatives. *Glycobiology,* 17**,** 983-993.

RUDD, T. R., UNIEWICZ, K. A., ORI, A., GUIMOND, S. E., SKIDMORE, M. A., GAUDESI, D., XU, R. Y., TURNBULL, J. E., GUERRINI, M., TORRI, G., SILIGARDI, G., WILKINSON, M. C., FERNIG, D. G. & YATES, E. A. 2010. Comparable stabilisation, structural changes and activities can be induced in FGF by a variety of HS and non-GAG analogues: implications for sequence-activity relationships. *Organic & Biomolecular Chemistry,* 8**,** 5390-5397.

RUDLAND, P. S., PLATTHIGGINS, A. M., WILKINSON, M. C. & FERNIG, D. G. 1993. Immunocytochemical identification of basic fibroblast growth factor in the developing rat mammary Gland variations in location are dependent on glandular structure and differentiation. *Journal of Histochemistry & Cytochemistry,* 41**,** 887-898.

RUDLAND, P. S., TWISTON DAVIES, A. C. & TSAO, S. W. 1984. Rat mammary preadipocytes in culture produce a trophic agent for mammary epithelia-prostaglandin E2. *J. Cell. Physiol.,* 120**,** 364-376.

SCHLESSINGER, J., PLOTNIKOV, A. N., IBRAHIMI, O. A., ELISEENKOVA, A. V., YEH, B. K., YAYON, A., LINHARDT, R. J. & MOHAMMADI, M. 2000. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Molecular Cell,* 6**,** 743-750.

SEMISOTNOV, G. V., RODIONOVA, N. A., RAZGULYAEV, O. I., UVERSKY, V. N., GRIPAS, A. F. & GILMANSHIN, R. I. 1991. Study of the molten globule intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers,* 31**,** 119-128.

SERLS, A. E., DOHERTY, S., PARVATIYAR, P., WELLS, J. M. & DEUTSCH, G. H. 2005. Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. *Development,* 132**,** 35-47.

SHUTE, J. K., SOLIC, N., SHIMIZU, J., MCCONNELL, W., REDINGTON, A. E. & HOWARTH, P. H. 2004. Epithelial expression and release of FGF-2 from heparan sulphate binding sites in bronchial tissue in asthma. *Thorax,* 59**,** 557-562.

SIMONS, K. & GERL, M. J. 2010. Revitalizing membrane rafts: new tools and insights. *Nature Reviews Molecular Cell Biology,* 11**,** 688-699.

SLEEMAN, M., FRASER, J., MCDONALD, M., YUAN, S. N., WHITE, D., GRANDISON, P., KUMBLE, K., WATSON, J. D. & MURISON, J. G.

2001. Identification of a new fibroblast growth factor receptor, FGFR5. *Gene*, 271**,** 171-182.

STATHOPOULOS, A., TAM, B., RONSHAUGEN, M., FRASCH, M. & LEVINE, M. 2004. pyramus and thisbe: FGF genes that pattern the mesoderm of Drosophila embryos. *Genes & Development,* 18**,** 687-699.

SUN, C., LI, Y., TAYLOR, S. E., MAO, X., WILKINSON, M. C. & FERNIG, D. G. 2015. HaloTag is an effective expression and solubilisation fusion partner for a range of fibroblast growth factors. *PeerJ,* 3**,** e1060.

SUTHERLAND, D., SAMAKOVLIS, C. & KRASNOW, M. A. 1996. Branchless encodes a Drosophila FGF homolog that controls tracheal cell migration and the pattern of branching. *Cell,* 87**,** 1091-1101.

SWEENEY, S. M., GUY, C. A., FIELDS, G. B. & SAN ANTONIO, J. D. 1998. Defining the domains of type I collagen involved in heparin-binding and endothelial tube formation. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 7275-7280.

TANAKA, K. A. K., SUZUKI, K. G. N., SHIRAI, Y. M., SHIBUTANI, S. T., MIYAHARA, M. S. H., TSUBOI, H., YAHARA, M., YOSHIMURA, A., MAYOR, S., FUJIWARA, T. K. & KUSUMI, A. 2010. Membrane molecules mobile even after chemical fixation. *Nature Methods,* 7**,** 865-866.

TAVERNA, S., GHERSI, G., GINESTRA, A., RIGOGLIUSO, S., PECORELLA, S., ALAIMO, G., SALADINO, F., DOLO, V., DELL'ERA, P., PAVAN, A., PIZZOLANTI, G., MIGNATTI, P., PRESTA, M. & VITTORELLI, M. L. 2003. Shedding of membrane vesicles mediates fibroblast growth factor-2 release from cells. *Journal of Biological Chemistry,* 278**,** 51911-51919.

TAYLOR, K. R. & GALLO, R. L. 2006. Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation. *Faseb Journal,* 20**,** 9-22.

TEKIN, M., HISMI, B. O., FITOZ, S., OZDAG, H., CENGIZ, F. B., SIRMACI, A., ASLAN, I., INCEOGLU, B., YUKSEL-KONUK, E. B., YILMAZ, S. T., YASUN, O. & AKAR, N. 2007. Homozygous mutations in fibroblast growth factor 3 are associated with a new form of syndromic deafness characterized by inner ear agenesis, microtia, and microdontia. *American Journal of Human Genetics,* 80**,** 338-344.

THEODORAKI, A., HU, Y., POOPALASUNDARAM, S., OOSTERHOF, A., GUIMOND, S. E., DISTERER, P., KHOO, B., HAUGE-EVANS, A. C., JONES, P. M., TURNBULL, J. E., VAN KUPPEVELT, T. H. & BOULOUX, P. M. 2015. Distinct patterns of heparan sulphate in pancreatic islets suggest novel roles in paracrine islet regulation. *Molecular and Cellular Endocrinology,* 399**,** 296-310.

THOMPSON, S. M., CONNELL, M. G., VAN KUPPEVELT, T. H., XU, R., TURNBULL, J. E., LOSTY, P. D., FERNIG, D. G. & JESUDASON, E. C. 2011. Structure and epitope distribution of heparan sulfate is disrupted in

experimental lung hypoplasia: a glycobiological epigenetic cause for malformation? *BMC Dev Biol,* 11**,** 1-17.

TUMOVA, S., WOODS, A. & COUCHMAN, J. R. 2000. Heparan sulfate proteoglycans on the cell surface: versatile coordinators of cellular functions. *International Journal of Biochemistry & Cell Biology,* 32**,** 269-288.

TURNBULL, J., POWELL, A. & GUIMOND, S. 2001. Heparan sulfate: decoding a dynamic multifunctional cell regulator. *Trends in Cell Biology,* 11**,** 75-82.

TURNBULL, J. E., FERNIG, D. G., KE, Y. Q., WILKINSON, M. C. & GALLAGHER, J. T. 1992. Identification of the basic fibroblast growth factor binding sequence in fibroblast heparan sulfate. *Journal of Biological Chemistry,* 267**,** 10337-10341.

TURNER, N. & GROSE, R. 2010. Fibroblast growth factor signalling: from development to cancer. *Nature Reviews Cancer,* 10**,** 116-129.

UNIEWICZ, K. A., ORI, A., XU, R. Y., AHMED, Y., WILKINSON, M. C., FERNIG, D. G. & YATES, E. A. 2010. Differential scanning fluorimetry measurement of protein stability changes upon binding to glycosaminoglycans: a screening test for binding specificity. *Analytical Chemistry,* 82**,** 3796-3802.

VIVOLI, M., NOVAK, H. R., LITTLECHILD, J. A. & HARMER, N. J. 2014. Determination of protein-ligand interactions using differential scanning fluorimetry. *Journal of Visualized Experiments***,** (91), e51809.

VLODAVSKY, I., FOLKMAN, J., SULLIVAN, R., FRIDMAN, R., ISHAIMICHAELI, R., SASSE, J. & KLAGSBRUN, M. 1987. Endothelial cell-derived basic fibroblast growth factor: synthesis and deposition into subendothelial extracellular matrix. *Proceedings of the National Academy of Sciences of the United States of America,* 84**,** 2292-2296.

WALKER, A., TURNBULL, J. E. & GALLAGHER, J. T. 1994. Specific heparan sulfate saccharides mediate the activity of basic fibroblast growth factor. *Journal of Biological Chemistry,* 269**,** 931-935.

WIEDEMANN, M. & TRUEB, B. 2000. Characterization of a novel protein (FGFRL1) from human cartilage related to FGF receptors. *Genomics,* 69**,** 275-279.

WRIGHT, P. E. & DYSON, H. J. 2015. Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology,* 16**,** 18-29.

XU, D. & ESKO, J. D. 2014. Demystifying heparan sulfate-protein interactions. *Annual Review of Biochemistry,* 83**,** 129-157.

XU, R. Y. 2012. *Analysis of fibroblast growth factor-heparin interactions (PhD Thesis).* PhD, University of Liverpool.

XU, R. Y., ORI, A., RUDD, T. R., UNIEWICZ, K. A., AHMED, Y. A., GUIMOND, S. E., SKIDMORE, M. A., SILIGARDI, G., YATES, E. A. & FERNIG, D. G. 2012. Diversification of the structural determinants of fibroblast growth factor-heparin interactions implications for binding specificity. *Journal of Biological Chemistry,* 287**,** 40061–40073.

XU, R. Y., RUDD, T. R., HUGHES, A. J., SILIGARDI, G., FERNIG, D. G. & YATES, E. A. 2013. Analysis of the fibroblast growth factor receptor (FGFR) signalling network with heparin as coreceptor: evidence for the expansion of the core FGFR signalling network. *Febs Journal,* 280**,** 2260-2270.

YAN, D. & LIN, X. H. 2009. Shaping morphogen gradients by proteoglycans. *Cold Spring Harbor Perspectives in Biology,* 1**,** a002493.

YANAGISHITA, M. & HASCALL, V. C. 1992. Cell surface heparan sulfate proteoglycans. *Journal of Biological Chemistry,* 267**,** 9451-9454.

YE, S., LUO, Y. D., LU, W. Q., JONES, R. B., LINHARDT, R. J., CAPILA, I., TOIDA, T., KAN, M., PELLETIER, H. & MCKEEHAN, W. L. 2001. Structural basis for interaction of FGF-1, FGF-2, and FGF-7 with different heparan sulfate motifs. *Biochemistry,* 40**,** 14429-14439.

YU, K. & ORNITZ, D. M. 2001. Uncoupling fibroblast growth factor receptor 2 ligand binding specificity leads to Apert syndrome-like phenotypes. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 3641-3643.

YU, S. R., BURKHARDT, M., NOWAK, M., RIES, J., PETRASEK, Z., SCHOLPP, S., SCHWILLE, P. & BRAND, M. 2009. Fgf8 morphogen gradient forms by a source-sink mechanism with freely diffusing molecules. *Nature,* 461**,** 533-539.

YUNG, S. & CHAN, T. M. 2007. Glycosaminoglycans and proteoglycans: Overlooked entities? *Peritoneal Dialysis International,* 27**,** S104-S109.

ZEHE, C., ENGLING, A., WEGEHINGEL, S., SCHAFER, T. & NICKEL, W. 2006. Cell-surface heparan sulfate proteoglycans are essential components of the unconventional export machinery of FGF-2. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 15479-15484.

ZELLER, R., LOPEZ-RIOS, J. & ZUNIGA, A. 2009. Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nature Reviews Genetics,* 10**,** 845-858.

ZHANG, D. Z., KOSMAN, J., CARMEAN, N., GRADY, R. & BASSUK, J. A. 2006a. FGF-10 and its receptor exhibit bidirectional paracrine targeting to urothelial and smooth muscle cells in the lower urinary tract. *American Journal of Physiology-Renal Physiology,* 291**,** F481-F494.

ZHANG, J. M., YANG, P. L. & GRAY, N. S. 2009. Targeting cancer with small molecule kinase inhibitors. *Nature Reviews Cancer,* 9**,** 28-39.

ZHANG, X. Q., IBRAHIMI, O. A., OLSEN, S. K., UMEMORI, H., MOHAMMADI, M. & ORNITZ, D. M. 2006b. Receptor specificity of the fibroblast growth factor family - The complete mammalian FGF family. *Journal of Biological Chemistry,* 281**,** 15694-15700.

ZHOU, S. H., LO, W. C., SUHALIM, J. L., DIGMAN, M. A., GRATTON, E., NIE, Q. & LANDER, A. D. 2012. Free extracellular diffusion creates the Dpp morphogen gradient of the drosophila wing disc. *Current Biology,* 22**,** 668-675.

ZHU, H. Y., DUCHESNE, L., RUDLAND, P. S. & FERNIG, D. G. 2010. The heparan sulfate co-receptor and the concentration of fibroblast growth factor-2 independently elicit different signalling patterns from the fibroblast growth factor receptor. *Cell Communication and Signaling,* 8**,** 14.

ZHU, X., KOMIYA, H., CHIRINO, A., FAHAM, S., FOX, G. M., ARAKAWA, T., HSU, B. T. & REES, D. C. 1991. Three-dimensional structures of acidic and basic fibroblast growth factors. *Science,* 251**,** 90-93.